

Closed caption, open source

E Blankinship, B Smith, H Holtzman and W Bender

This paper explores how television fans appropriate video for personal expression and how technology can support such creative appropriation. Televisions do not have an equivalent to a Web browser's view source option; however, programmes can be structured by their transcripts, embedded as closed captions in the signal of most shows. With our talkTV video-editing software, rearranging lines of dialogue automatically creates new scenes, thereby enabling television viewers to become authors and editors. We present a case study of how television fans used talkTV and conclude with a discussion of the implications to digital rights management of this work and the potential for 'view source' television.

1. Introduction

Television is often considered to be a passive medium, as viewers generally sit back and absorb whatever content is being played [1]. Yet there are some television and movie fans who are inspired to play a more active role; they re-edit broadcast television shows, movies, and clips from the Internet for creative expression. For instance, a video clip of a hapless kid pretending to be a character from the Star Wars movies turned up on the Internet and shortly afterwards many amateur video editors downloaded that clip, added visual effects and soundtracks, then shared their products with others on-line (see Fig 1) [2].

Our research focuses on these creative individuals, investigating ways to index and deliver video to them for reuse. We have created a video editing system, talkTV, which uses closed-caption subtitle tracks from television programmes to be searched and manipulated into novel presentations. Our hope is that by providing better ways to

organise television for creative projects and an argument for the need for better ways, passive television viewing will be supplemented by constructive television production.

This paper begins with a short history of how television fans and artists have re-edited television shows for creative projects and the need for better ways to index televised content. The next section describes how closed-captioning has been used as an index for computers to search and manipulate television programmes such as news broadcasts. Then our talkTV software is described as a special-purpose video editor that television fans can use to re-sequence closed-captioned video clips. A case study of how television fans used talkTV describes how our video editor supports creative appropriation of television. As the sampling of media from various copyright owners has inherent legal issues, we describe a potential solution for content reuse. In conclusion, we advocate that television programmes and other media be accessible via mark-up languages, which offer semantic descriptions of the



Fig 1 Screen shots from different fan versions of the Star Wars Kid video in which a kid practices dueling with a toy sword. On the left is *The Matrix Star Wars Kid* in which background footage is composited from the movie *The Matrix*. In the middle is *Animated Star Wars Kid*, a rotoscoped version of the video. On the right is *Tetris Star Wars Kid* in which he knocks out falling video game blocks. More fan versions of the *Star Wars* kid are available at <http://www.jedimaster.net/>

content, so that their presentation can be more easily manipulated for creative projects.

2. Creative appropriation of television

In the late 1970s, television fans started re-editing television shows for personal expression. By connecting the video output from one consumer VCR to the video input on another VCR, these fans dubbed cued video clips. In this way, television fans assembled their own montages, often set to popular music. Screenings of fan vids took place at science-fiction conventions, where other fans learned the procedure, thus proliferating the art of fan viding, which remains popular today. A well-crafted fan vid juxtaposes video clips with relevant song lyrics; thereby creating a novel interpretation of the television show's canon. For example, a well produced Star Trek fan vid juxtaposes video clips of Kirk and Spock gazing into the distance cued to the lyrics of a romantic ballad, thereby suggesting romance between these characters [3]. These vids assume an audience not only familiar with the show and its characters, but also with meanings and vocabulary of the fan culture that produces them. For example, the fan genre of slash fiction (in written works) had been established for a decade before vids were produced.

To find desired video clips, fans must have an exhaustive knowledge of a television show. To help find video clips near memorable events, some fans produce 'clip notes', which are listings of approximate times for memorable events in certain episodes. Fan vidders are not the only ones indexing content for creative repurposing; popular artists are also combing through hours of television:

Working with a computer database, [artists Jennifer and Kevin McCoy] have sliced and diced several emblematic television shows into their component parts. A result of this archaeology of entertainment is the digital equivalent of thousands of carefully labelled pot shards.

The McCoys sorted and categorised scenes from *Kung Fu* and postwar *Looney Tunes* on DVDs, which expose the viewer to endless variations on blocked punches or explosions. They dissected an episode from *Eight Is Enough* into 448 shots, also on DVDs, arranged in a big colour-coordinated grid on the wall. They spliced together a disjunctive show-within-a-show using all the action shown on the big flight-deck screen of the Starship Enterprise from the first season of *Star Trek* [4].

For the McCoys, creating indices with which to disassemble and then reassemble television shows is creative interpretation and commentary on the shows' content. Sophomoric humourists at MAD Magazine [5] have also repurposed television to this effect, noting: '[television] programs are as boring as ever, so now [people] spend all of their time pushing the button on the cable box to switch from station to station, which creates a strange kind of craziness we call... MAD'S CABLE TV ROULETTE.'

MAD's Roulette is a comic storyboard of frames grabbed from various television programmes in which television hosts complete each other's sentences, such as Chef Julia Child's

instructions for a recipe being completed by Dr Ruth, sex therapist.

Although the comic dialogue in this storyboard is scripted, it demonstrates how dialogue can be used for creative repurposing.

Dialogue is a useful index into video, as people tend to remember dialogue. In fact, people regularly appropriate television dialogue when they cite familiar quotes in their conversations. Quips such as 'Live long and prosper' (Star Trek), 'D'oh!' (The Simpsons), and 'Wassup!' (Budweiser commercial) convey a shared meaning between people. Appropriated television and movie quotes are also common in e-mail signature lines, often selected for a personal flourish. Another use of dialogue for conversational purposes, albeit primarily prank calls, is via soundboard software, which plays celebrity sound bites at the push of a button. Soundboard creators search through television shows for audio clips to embed in their soundboards, much like fan vidders search for video clips.

in the late 1970s, television fans started re-editing television shows for personal expression

Searching through video for a specific clip is not an easy task, and consumer devices such as VCRs only offer rewinding, pausing, and fast forwarding as their main search features. Perpetuating television as a browse only medium, DVDs structure their contents in the form of video chapters within a menu structure. Personal video recorders (PVRs) such as TiVo also list recorded shows in menus. The next section of this paper describes how closed-captioned dialogue has been used as an alternative index into television.

3. Alpha bits — indexing television with subtitles

With a pioneering spirit, and a bit of hacking, computer researchers in the 1980s began to play with television signals to see how they could be digitally manipulated. Closed-captioning, used primarily to broadcast subtitles on the vertical blanking interval of a television signal, presented itself as a useful index into televised data. Because of legislation [6], closed captions are broadcast with a wide and expanding variety of programming.

Searching through digitised text is less computationally expensive than searching through image data, and has the additional benefit of being relatively easy to query. One early project that used computational search of television was Network Plus [7], which augmented televised news broadcasts with contextual information from the Associated Press news wire. Keywords extracted from the closed-caption data were used to retrieve relevant information from external databases, such as the wire services, and this information was presented as a synchronous overlay to the broadcast programme using

the closed captioning as a time code. Another project, Auto Cassette [8], parsed the closed captioning from evening news broadcasts to automatically segment news stories and create a table of contents of the broadcast. Titles for the news stories were extracted from the first sentence following the newscaster's introduction.

Despite closed captioning offering both a temporal and descriptive index into televised content, most consumer television devices today have relegated closed captions to be used as a display-only mechanism. However, two consumer devices use the closed captions to prevent the display of subtitled dialogue — Curse-Free TV™ and TVGuardian™ scan closed captions for profanity and both silence corresponding offensive audio and present alternate subtitles. Some personal-computer video-capture cards also scan closed captions, and have bundled software to alert users when key words are spotted or to create written transcripts (without time code). Consumer devices treat subtitles as just another track, which can be turned on and off like video or audio, instead of using the information to organise television content.

talkTV is a graphic editor designed to assist television fans re-sequence television dialogue

The advent of the World Wide Web popularised a new way of using text to organise multimedia. Web pages are written as text files using hypertext markup language (HTML) to structure their hyperlinks, graphical layout, and multimedia content for a Web browser to display. Every Web browser since the original Mosaic browser has a 'view source' menu item that reveals these HTML source files, making it possible for people to copy and modify a Web page and use it as a template for their own page [9]. Television, a far more ubiquitous medium, does not have an equivalent option allowing people to 'look under the bonnet'.

Could television shows be organised like Web pages so as to be more easily modified? Closed captions can be saved as a text file that references and sequences many multimedia snippets of dialogue. Just as a Web page's HTML code may reference other files such as GIF and JPEG images, each line of subtitled dialogue references a snippet of audio and video. However, televisions do not use closed captions to structure the display of television programmes in the same way that a Web browser uses HTML. But if television shows were organised this way, reordering the sequence of the closed captions would create new videos. With this in mind, we created a video editor to treat closed captioning as if it were a markup language to see how it would effect television fans' abilities to reuse television content in creative projects.

4. From remote control to direct control

This section of the paper describes how our talkTV software assists television fans with their re-editing of television shows.

Television shows used for this research project were digitised from broadcasts, VHS tapes, and DVDs into MPEG1 movie files of their video and audio content and XML text files of their closed captions (see Fig 2). Digitised television shows are placed into a folder where talkTV loads in their closed-captioned transcript, which includes time code of when a line of dialogue approximately begins and ends in the television show. The talkTV editor segments television shows into small clips of dialogue by marking in points at the beginning of a closed caption and marking out points on stop characters in the subtitles such as periods, question marks, and exclamation points. In this way, talkTV processing roughly parses the content at the boundaries where dialogue exchanges take place.

```
<?xml version='1.0' encoding='UTF-8'?>
- <EPISODE mov='101.mov'>
- <clip num='0'>
  <source line='Space, the final frontier.'
    tc='00:00:20.11' tcout='00:00:25.25' />
  <mods />
</clip>
- <clip num='1'>
  <source line='These are the voyages of the
    Starship Enterprise.'
    tc='00:00:25.25' tcout='00:00:29.20' />
  <mods />
</clip>
- <clip num='2'>
  <source line='Its continuing mission...'
    tc='00:00:29.21' tcout='00:00:31.20' />
  <mods />
</clip>
```

Fig 2 A sample of our XML representation of closed caption data. Time code in points and out points are indicated by attribute elements such as 'tcin=' and 'tcout=' to mark the boundaries of dialogue exchanges. Adjustments to these times can be made through the talkTV application; any modifications are saved to the 'mods' element so that the original time code is not lost.

TalkTV is a graphic editor designed to assist television fans re-sequence television dialogue. Since talkTV uses closed captions in the form of text documents to organise television content, a text-only document editor could also be custom built wherein time-coded subtitle elements from various closed caption XML documents could be cut and pasted together to create new fan videos. However, talkTV was built in the spirit of Web page editors such as Netscape Composer™, Microsoft FrontPage™, and Macromedia Dreamweaver™, which assist in the modifying of HTML files so that Web page layouts can be graphically manipulated, helping to ensure that 'what you see is what you get'.

4.1 Searching for dialogue

There are several ways television fans can use talkTV to find television dialogue. One way is to search for dialogue by typing words into a text field. For example, a fan might type in 'go out' to retrieve all of the video clips containing these words in the spoken dialogue and the closed captions (see Fig 3, upper left). Video clip search results are presented as a list of thumbnails that can be played to see and hear the dialogue as performed. The complete text of the clip's closed-captioned dialogue (e.g. 'Don't you ever go out with anyone else at all?') is displayed as text under its thumbnail for quick visual browsing (Fig 3, lower left).

Another way a television fan can find dialogue clips is by browsing through an alphabetical list of all of the words in every transcript. When a word is selected from the list, clips of dialogue with that word in their closed captions are retrieved and displayed as search results (Fig 3, upper left).

Television fans can also browse through the dialogue of an individual television show to find clips. Using the 'episode view' panel, a television fan can select one television show from a menu to display its dialogue exchanges in a sequential list. Fans can select a line of dialogue from the list to display its video clip.

If the beginning and ending of the clip do not precisely match the subtitled text, television fans can fix the discrepancy by using trimming options under the 'clip' menu, allowing an additional half a second to be added or removed to the beginning or ending of a clip. Since closed-captioned dialogue is added as one of the final steps in the production cycle for a pre-produced programme, it provides only a modicum of accuracy. Live broadcast captions are often rife with misspellings and are by rule delayed by up to several seconds, but these can be time shifted by talkTV to account for the transcription delays.

4.2 Copying a dialogue video clip into a fan video

TalkTV has a graphic storyboard panel where television dialogue is re-sequenced. The storyboard panel is a grid of twenty-five individual storyboards (Fig 3, right). Using the mouse, television fans can drag dialogue video clips from their search results and drop the clips on to an individual storyboard. Dropping a video clip on to a storyboard makes a copy of the video clip and also compresses the video clip into a

smaller file. Once copied into the storyboard, video clips can be re-sequenced by dragging them to another storyboard, or deleted by pressing a storyboard's 'delete' button.

By pressing the 'play' button, video clips in the storyboard are edited together into a new video file and played as one movie in a full-screen preview. Playing re-sequenced video clips edited together into a new movie is talkTV's equivalent to reloading an edited Web page in a Web browser.

4.3 Adding titles to a fan video

Television fans can use talkTV to create and insert titles (or interstitial text) into their projects. Television fans can do this by typing title text into a text field, and then previewing their text superimposed over pre-selected background videos. Title video clips can be dragged and dropped into the storyboard just like dialogue video clips. This tool was intended to create inter-titles in the style of a silent film or a coming attraction, so as to help with continuity between different dialogue video clips.

4.4 Recording new video for the storyboard

With an attached Web camera, fans can record short videos to drag into the storyboard. Fans can use this feature to personalise a vid by inserting clips of themselves acting, establishing shots, props, or other events.

4.5 Sharing your fan video electronically

Fan videos made with talkTV can be saved as a Quicktime™ digital-video file. If a fan video is e-mailed to another television fan, they can watch the video. If the other fan also has the talkTV application, they can import the video into their own



Fig 3 A screen capture of the talkTV software annotated with numbers for instructions: (1) search for dialogue by typing your search query into the text field; (2) browse dialogue video clips by clicking on the play button under each video clip to watch that video clip; and (3) drag dialogue search results to the storyboard panel where they are re-sequenced.

collection of television shows. This is because talkTV embeds the original closed-captioned source code into every fan video. After a fan video is imported, it can be searched for dialogue clips and those clips can be incorporated into new projects.

To make it easy to send fan videos to other television fans, talkTV has a graphic 'drag strip', where the edited-together version of the fan video in the storyboard panel is ready for dragging into other applications like e-mail and chat. Individual storyboard clips can also be dragged and copied into other applications.

5. To boldly go — sci-fi fans make videos with talkTV

In order to observe how people might use talkTV, we brought it to science-fiction fans at an annual fan run convention in Massachusetts. Two computer terminals running talkTV shared access to over 100 hours of Star Trek, and completed fan videos were played on a large, publicly viewable screen. In total, 21 fan videos were created. Because of concern over copyright, fans were not able to electronically send their videos to anyone, nor were they allowed to keep a copy. In response to a survey, we learned that 90% of the participants had never edited film or video before using talkTV, but all were familiar with fan vids. Most said fan viding was a specialised skill requiring a lot of time, but that they would like to try it some time.

'my own editing lab at my fingertips'

The majority of participants endeavoured to create mini-episodes or stereotypical scenes from the Star Trek series. For example, some fans would begin using talkTV by searching for words like 'enemy ships approaching', browse through the video-clip search results, and drag one to the storyboard. They would continue in this fashion, sequentially adding clips to the storyboard with dialogue such as 'we're under attack', 'shields up', 'return fire', and 'retreat'. After collecting all of these clips into their storyboard, fans would satisfactorily preview their completed fan video.

While there was not always a strong visual continuity between the clips, other fans of the series could follow and appreciate the drama when the video was played on the public monitor.

Other fans searched for a specific word, like 'Klingon', and happily filled their entire storyboard panel with dialogue about Klingons. No apparent visual continuity applied to these fan videos, which played like highlight reels from the television series.

A few fans endeavoured to create scenes with dialogue exchanges throughout. Many of these fans started their videos by finding dialogue of a provocative question, such as 'Where are we?' or 'What is going on?' Once this question was dragged to the storyboard panel, they would browse through many video clips to find a video clip with a reply. Videos made in this fashion often used repetition of a question with varied responses for humorous effect.

Finding video clips that fit into their scene was described as analogous to completing a puzzle (Fig 4). In 'finding a way to tie it together', the participants were proud of their ability to effectively mine the dialogue to create their videos. One user described it as a 'problem-solving' challenge, while another commented, 'It was funny to select [clips] and make them fit... [and] taking them out of context.' The participants often described their sense of accomplishment in terms of feeling like an actual director or screenwriter, occupations that affirm their appropriation of the media.

Several of the users had memorised whole seasons of the television show. One fan commented: 'It's a good way of testing your knowledge of the show. It triggers recall of episodes.' Their ability to remember the scripts of various scenes enabled them to query efficiently for specific characters or locales. Less familiar fans still found the query searches 'fun' because they were able to re-watch the television clips.

It took a while for the participants to understand that only closed captions were used to organise the available video clips. For example, a few participants queried for 'explosion' and expressed dissatisfaction when only dialogue about explosions was retrieved. Once participants understood the limitations of the index, they were able to quickly assemble their videos as described above. As one participant said, '[It's] fun to query clips and then you can cut and paste ... my own editing lab at my fingertips'.

6. Whose line is it anyway?

Copyright law establishes a legal framework for ownership of creative expression that grants exclusive control of copyrightable material to the owner. Because talkTV provides a system for creating new movies composed of snippets of existing works, it would be remiss to not say a few words about distributing the resulting works. Each of the original works that are used to create talkTV output is likely owned by some party and under their control. Reusing those works could entail securing permission from the owners of the original works.

Securing rights is an onerous task¹. The ordinary, though creative, individual — such as the fan who is the target user of

¹ 'When in doubt as to the right to use or acknowledge a source, the most prudent course may be to seek permission of the copyright owner' [11]. However, investing the time and effort to secure permission to use copyright material is no guarantee that permission will be granted. The authors were unable to secure rights from Viacom to use frames 'grabbed' from *Star Trek: The Next Generation* ... to illustrate talkTV '... due to a variety of legal and business considerations.' We were also unable to secure rights from New Line Productions to use a single frame from the movie, *Lord of the Rings: Fellowship of the Ring*, to illustrate adaptation: '... the reason has to do with [New Line Production's] relationship with the Tolkien estate.' Ironically, DC Comics was 'uncomfortable' granting permission to use an image from MAD Magazine's Cable TV Roulette since the use of talkTV 'requires the downloading and/or copying of copyrighted works that may be deemed infringing.' They added that they '... generally do not review academic articles but rather refer authors to check with their legal counsel to review the potential fair use nature of the use — unfortunately we do not have a large enough staff to review academic articles that use our materials, and forego the review of such articles since most of them appear, at least on first glance and limited review, to have good fair use arguments.' However, '... the only way to get a definitive answer on whether a particular use is a fair use is to have it resolved in federal court.' [10].

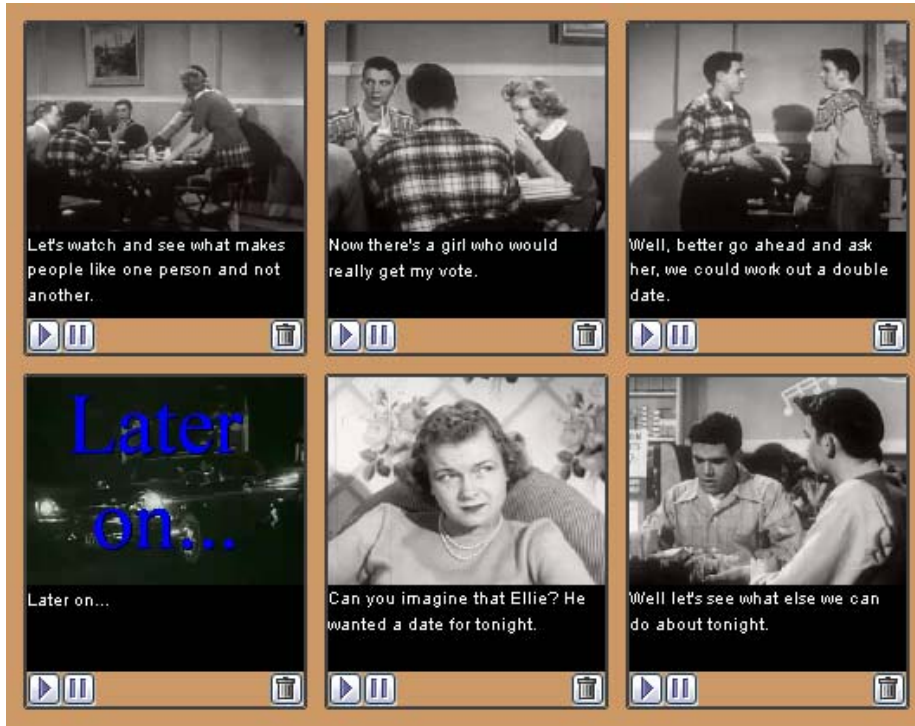


Fig 4 An example of a storyboard made with the talkTV video editing application. The vid begins with introductory narration, followed by video clips in which students discuss asking girls out for dates. The intertitle 'Later on ...' is followed by video clips in which the resloution of their proposal is discussed. Dialogue video clips are from the film 'Are You Popular?' [11].

talkTV — has no profit to share with the owners of the original works, dooming such negotiations from the start. TV and movie producers spend extraordinary resources to deal with this same issue for their own productions because much of what appears within the frame of a TV show or movie can be claimed subject to copyright law, whether it is a piece of designer furniture, a statue, or other work of art, and the owners of those copyrights must grant permission for the eventual distribution of the movie or TV show. As high budget releases have been held ransom to such negotiations, film makers now deal with this issue right from the start of production, trying to keep any copyrighted work out of the shot unless permission to use it has been pre-arranged.

While copyright law varies from country to country and enforcement from court to court, there are two common concepts that play into these reuse scenarios — derivative works and fair use. A derivative work is one that takes an existing work and creatively transforms it. The copyright owner of one of the original works incorporated into a talkTV movie might claim that the talkTV movie is a derivative work of their original. Generally, copyright law gives control of the creation of derivative works to the owner of the copyright on the original work. This control is what prevents authors from creating their own, unapproved 'Star Wars' novels and movies, for example.

Striking a balance on the side of free speech is 'fair use'. Fair use recognises that criticism, comment, news reporting, teaching, scholarship, and research are all important societal endeavours. In the USA, adjudicating fair use is a complex matter where purpose and character of the use (e.g.

commercial versus non-commercial), the nature of the copyrighted work, the amount of the whole work that is used, and the effect on the market for the original work are all considered.

Perhaps the issue of copyright and derivative works can be skirted by creating a distribution means that does not distribute a derivative work *per se*. Instead of distributing a new movie created by talkTV, we could distribute instructions for creating that movie. If the recipient of the instructions has legally obtained the right to view the original works, for example by purchasing copies of the original works, then copyright law may look favourably on the recipient of those instructions reforming their copies in creative ways for their own enjoyment. Permission may not be needed in such a scenario.

In the early 1990s, researchers at the MIT Media Lab envisioned a multimedia distribution infrastructure (the Media Bank) that would promote content reuse in such a fashion [12]. The Media Bank was an early peer-to-peer content delivery network. Every node of the Media Bank is both a server as well as a consumer of content. Content is replicated on multiple nodes using the Media Bank, providing increased availability, both in terms of fault tolerance and download bandwidth. Playback of a multimedia presentation can involve simultaneous download from multiple servers.

To honour copyright constraints, the Media Bank does not implicitly replicate content among its nodes, unlike systems such as Kazaa and Bit Torrent. Content is explicitly replicated on servers that have permission to offer that content for

distribution. Another key differentiating attribute of the Media Bank is that content is segmented into content-meaningful chunks. For example, a movie might be broken up for transfer into its constituent scenes, or line-by-line. These content-meaningful segments become the building blocks for the complete work, and it is by using these segments, rather than some arbitrary data chunking, that fault tolerance and bandwidth gains are achieved. Segmentation is desirable because multimedia content is often encoded using file formats and compression schemes that require the entire file be reassembled for playback to occur without errors.

two common concepts are at play — derivative works and fair use

Along with the content-based segmentation the Media Bank employs a scripting language for reassembling the movie at the receiver. The scripting language has primitives for specifying where Media Bank objects fit into the playback timeline. Multiple objects can be ‘in play’ at the same time, synchronised at the receiver. For example, this allows for storing soundtracks separately from the corresponding video. It is envisioned that third-party services could thereby come into existence, such as translating subtitles into new languages.

Content reuse is achieved using the Media Bank by distributing scripts that refer to the original content objects by name rather than incorporating their contents into new objects. The scripting language supports playing back subsections of objects, as well as assembling objects on the timeline in any order.

Objects from multiple servers and multiple original works can be arbitrarily assembled under the receiver’s control. Thus the receiver fetches the original content objects from duly authorised servers. If there are any access rights to be established, or fees to be paid this is negotiated directly between the server hosting the original content objects and the receiver. The creator of the script does not have to be in the middle of that negotiation.

7. On next week’s episode

We chose to observe science fiction fans using our software because these fans are easy to find *en masse* at conventions and have a history of media appropriation. However, there is little reason to think that these are the only types of television fan who will enjoy this sort of activity. Thousands of hours of television and movies in every genre are captioned and ready for their fans to play with. Being constructive with media is a good thing, as it extends media literacy to include authorship, and editors, like talkTV, make it easier for any television fan to be constructive with what’s playing on television.

7.1 *You can’t do that on television (yet)!*

Just as the World Wide Web provided fertile ground for many creative projects, we imagine that ‘view source’ television might inspire other projects. For example, many books are

adapted into movies and sometimes share the same dialogue. We linked the closed captions in the popular movie *Lord of the Rings: Fellowship of the Ring* to matching dialogue in the original book, and hand-marked-up additional text passages by noting the time-code of these scenes in the movie. Our software program, Adaptation, displays linkages between the source material and the movie. It does this by displaying the book and the movie as two parallel timelines, and graphically connecting corresponding sections to show how the book may have been expanded or condensed in adaptation.

Adaptation also calculates how many ‘minutes per page’ are spent on adapted scenes, and contrasts the relative size of scenes in their respective media.

7.2 *But wait, there’s more!*

Extensible markup language (XML) is a superset of the World Wide Web’s HTML for describing arbitrary data types. We advocate that television programmes be distributed as XML documents that organise television content. Likewise, we advocate that televisions act as extensible stylesheet language transformation (XSLT) processors. XSLT was ‘... originally intended to perform complex styling operations, like the generation of tables of contents and indexes’, but as we suggest with these closing examples, there are many creative ways to render organised sequential multimedia.

- New wireless telephones can display short video messages. Dialogue video clips can be sent as MMS messages to these mobile telephones. This application augments established popular SMS services by allowing television celebrities to present short messages. The simple talkTV interface of search and sequence is readily modifiable to support the small-screen constraints of a mobile telephone.
- While working on text documents such as e-mail, typed-in phrases that match closed-captioned phrases can be automatically replaced with video clips in that document.
- Taking our cue from some World Wide Web proxy servers that replace keywords and alter graphics on Web pages as you browse, we propose television proxy servers. By scanning incoming closed captions for dialogue also in other television programmes, any video of dialogue matches could be displayed inside a picture-in-a-picture.
- Similarly, Web pages could be run through television proxy servers, replacing text on Web pages with small video clips with dialogue matches. In this way, you could surf the Web with clips of your favourite television programme.
- Text-to-speech software can analyse children’s chatter when they play with toys, and spoken phrases matching captioned phrases can play those audio clips aloud [13]. The interaction would be akin to improvisational theatre, in which audience members throw out suggestions to the actors to incorporate into their performance.
- Television dubbed in multiple languages can automatically alternate the spoken language at every dialogue exchange.

Closed caption, open source

- Even the dullest PowerPoint presentation can be livened up with interstitial comments by your favourite television stars.
- Similar to how talkTV indexes television shows with subtitles, karaoke song lyrics are time coded and could also be used as an index for similar projects to those described here.

Sensor data can also be used to mark up events in a television programme. For example, we built a system in which a kitchen pitcher was rigged with a water sensor and temperature sensor. These sensors transmitted their values via infra-red to a computer where they were time stamped and recorded to a sensor track synchronised with the video and audio track. We used this system to record a cooking show in which we prepare iced tea. The tea recipe is simple: hot water is poured into the pitcher, ice cubes are added to cool down the tea, and then tall glasses of iced tea are poured out. Sensor equipped pitchers can then be used for search queries into this video. For example, when the pitcher is filled with hot water, the video jumps to that scene in the video. If cool water is placed in the pitcher, the video jumps to the scene just after the ice cubes were added. If the pitcher is emptied, two video clip search results are returned — one of the iced tea being poured out, and another from the beginning of the show (both of these clips have the same sensor values: room temperature and no water). Sensor data is continually streamed to the computer. This means that search queries are nearly continuous without the need to specify a new query. In this way, you can work in the kitchen freely and video is retrieved in a 'just in time' fashion.

As this example suggests, closed captions are by no means the only document structure that can organise televised content. New industry initiatives to describe key visual elements in television, such as Descriptive Video Service (DVS) [14], could be used in similar ways. The cable television adage, 'a thousand channels and nothing on' might soon be remedied when television fans edit together the next thousand channels.

Acknowledgements

The authors would like to thank Levar Burton and Gates McFadden for many hours of inspiration and entertainment. We would also like to thank the members of the Media Lab's information:organized / Simplicity consortium, specifically Dr Marko Turpeinen and Alma Media, for their generous support of this work.

References

- 1 Postman N: 'Amusing ourselves to death: public discourse in the age of show business', New York, Penguin Books (1985).
- 2 Harmon A: 'Compressed data; fame is no laughing matter for the *Star Wars Kid*', The New York Times, New York (2003).
- 3 Jenkins H: 'Textual poachers: television fans and participatory culture', New York, Routledge (1992).
- 4 Smith R: 'Art in review — Jennifer and Kevin McCoy', The New York Times, New York (2002).
- 5 Snider M: 'MAD's cable TV roulette', MAD, pp 10—12 (1986).
- 6 Telecommunications Act of 1996, Pub LA No 104—04, 110 Stat 56: 'What is XSL?', retrieved July 2004 — <http://www.w3.org/Style/XSL/WhatIsXSL.html>
- 7 Bender W and Chesnais P: 'Network Plus', Proceedings of the SPIE Electronic Image Devices and Systems Symposium, Los Angeles, CA (January 1988).
- 8 Konishi A: 'Auto-Cassette: The Automatic Clipping Service For TV News', Electrical Engineering and Computer Science, Cambridge, Massachusetts Institute of Technology (1986).
- 9 O'Reilly T: 'Where the Web leads us', O'Reilly XML.com (1999) retrieved July 2004 — <http://www.xml.com/pub/a/1999/10/tokyo.html>
- 10 Copyright — http://fairuse.stanford.edu/Copyright_and_Fair_Use_Overview/chapter9/9-b.html
- 11 Coronet Instructional Films: 'Are You Popular', (1947) — <http://www.archive.org/movies/movies-details-db.php?collection=prelinger&collectionid=00014>
- 12 Lippman A and Kermode R: 'Media Banks: Entertainment and the Internet', IBM Systems Journal, 35, Nos 3 and 4, pp 272—291 (1996).
- 13 Sjursen E and Blankinship E: 'TVplaySET', (July 2004) — <http://web.media.mit.edu/~erikb/talkTV/playset/TVplayset.pdf>
- 14 WGBH: 'Media Access Group', DVS FAQ, retrieved July 2004 — <http://main.wgbh.org/wgbh/pages/mag/services/description/dvs-faq.html>



Erik Blankinship is developing a children's construction kit for designing computer opponents. If children explicitly described their implicit strategies in games, they might improve their understanding of mathematics.

A PhD candidate, he received an MS in media arts and sciences from MIT, an MEd from Harvard University, and a BA with a concentration in folklore and mythology from the University of Maryland at College Park.



Brian Smith is associate professor of information sciences and technology and associate professor of instructional systems in the School of Information Sciences and Technology at Penn State University. His research deals with computer-aided cognition and decision-making — how computational artefacts can be designed to enhance human learning and performance in 'everyday' settings. Prior to joining Penn State, he was on the faculty of the MIT Media Lab. He was featured in a special issue of *Black Issues in Higher Education* as one of ten

influential African-American innovators in information technology in February 2002.



Affiliated with MIT since 1981, Henry Holtzman is a research scientist at the MIT Media Lab examining tangible networking and image compression.

He is an expert in radio frequency identification (RFID), and represented MIT and the United States as a member of the MPEG standardisation committee, which helped to define the underlying video technology used in DirectTV, DVD, digital cable, and digital TV broadcasting.

He earned a BS in computer science and an MS degree in media arts and sciences, both from MIT.



Walter Bender is executive director of the MIT Media Laboratory, a senior research scientist, director of the Electronic Publishing group, and a member of the Laboratory's Information Organised consortium. He also directs the Gray Matters special interest group, which focuses on technology's impact on the aging population. A founding member of the Media Lab, he studies new information technologies, building upon the interactive styles associated with existing media and extending them into domains where a computer is incorporated into the interaction. He has participated in much of the pioneering research in the field of electronic publishing and personalised interactive multimedia. He received his BA from Harvard University, and his MS from MIT.