MUMIS – A Multimedia Indexing and Searching Environment

Declerck Thierry and Peter Wittenburg
DFKI GmbH and Max Planck Institute for Psycholinguistics
Stuhlsatzenhausweg 3, D-66123 Saarbruecken, Germany and
Wundtlaan 1,PB 310, 6500 AH Nijmegen, The Netherlands
declerck@dfki.de and Peter.Wittenburg@mpi.nl

We describe in this paper the MUMIS Project (Multimedia Indexing and Searching Environment)[1] and show the role linguistically motivated annotations, coupled with domain-specific information, can play for the indexing and the searching of multimedia (and multilingual) data. MUMIS develops and integrates base technologies, demonstrated within a laboratory prototype, to support automated multimedia indexing and to facilitate search and retrieval from multimedia databases. The project will demonstrate that innovative technology components can operate on multilingual, multisource, and multimodal information and create a meaningful and queryable database.

1. Goals of the MUMIS Project

MUMIS develops and integrates basic technologies, which will be demonstrated within a laboratory prototype, for the automatic indexing of multimedia programme material. Various technology components operating offline will generate formal annotations of events in the data material processed. These formal annotations will form the basis for the integral online part of the MUMIS project, consisting of a user interface allowing the querying of videos. The indexing of the video material with relevant events will be done along the line of time codes extracted from the various documents.

For this purpose the project makes use of data from different media sources (textual documents, radio and television broadcasts) to build a specialized set of lexicons and an ontology for the selected domain (soccer). It also digitizes non-text data and applies speech recognition techniques to extract text for the purpose of annotation.

The core linguistic processing for the annotation of the multimedia material consists of advanced information extraction techniques for identifying, collecting and normalizing significant text elements (such as the names of players in a team, goals scored, time points or sequences etc.) which are critical for the appropriate annotation of the multimedia material in the case of soccer.

Due to the fact that the project is accessing and processing distinct media in distinct languages, there is a need for a novel type of merging tool in order to combine the semantically related annotations generated from those different data sources, and to detect inconsistencies and/or redundancies within the combined annotations. The merged annotations will be stored in a database, where they will be combined with relevant metadata.

Finally the project will develop a user interface (in Dutch, English, German and Swedish) to enable professional users to query the database, by selecting from menus based on annotations, metadata, and so on, and to view video fragments retrieved to satisfy the query, offering thus a multimedia tool to formulate multilingual queries about multimedia programmes and directly get interactive access to the multimedia contents. This tool constitutes the online component of the MUMIS environment.

2. State of the Art

MUMIS differs in many significant ways from existing technologies and projects already funded. Most closely related to the thematic focus of MUMIS are the HLT projects Pop-Eye [POP] and OLIVE [OLI]. Pop-Eye used subtitles to index video streams and offered time-stamped texts to satisfy a user query, on request displaying a storyboard or video fragment corresponding to the text hit. OLIVE

[1]MUMIS is an EU-funded project within the Information Society Programme (IST) of the European Commission, section Human Language Technology (HLT). MUMIS had its kick-off meeting in September 2000.

used automatic speech recognition to generate transcriptions of the sound tracks of news reports, which were then indexed and used in ways similar to the Pop-Eye project; both projects used fuzzy matching IR algorithms to search and retrieve text, offering limited multilingual access to texts. Instead of using IR methods to index and search the transcriptions, MUMIS will create formal annotations to the information, and will fuse information annotations from different media sources. The fusion result is then used to direct retrieval, through interface techniques such as pop-up menus, keyword lists, and so on. Search takes the user direct to the storyboard and video clippings.

The Informedia project at Carnegie-Mellon-University [INF] has a similar conceptual baseline to MUMIS. The innovative contribution of MUMIS is that it uses a variety of multilingual information sources and fuses them on the basis of formal domain-specific annotations. Where Informedia primarily focuses on special applications, MUMIS aims at the advancement and integratibility of HLT-enhanced modules to enable information filtering beyond the textual domain. Therefore, MUMIS can be seen as complementary to Informedia with extensions typical for Europe.

The THISL project [THI] is about spoken document retrieval, i.e., automatic speech recognition is used to auto-transcribe news reports and then information retrieval is carried out on this information. One main focus of THISL is to improve speech recognition. Compared to MUMIS it lacks the strong language processing aspects, the fusion of multilingual sources, and the multimedia delivery.

Columbia university is running a project [COL] to use textual annotations of video streams to indicate moments of interest, in order to limit the scope of the video processing task which requires extreme CPU capacities. So the focus is on finding strategies to limit video processing. The University of Massachusetts (Amherst) is also running projects about video indexing [UMA], but these focus on the combination of text and images. Associated text is used to facilitate indexing of video content. Both projects are funded under the NSF Stimulate programme [NSF].

Much work has been done on video and image processing (Virage [VIR], the EUROMEDIA project [EUR], Surfimage [SUR], the ISIS project [ISI], IBM's Media Miner, projects funded under the NSF Stimulate program [NSF], and many others). Although this technology in general is in its infancy, there is reliable technology to indicate, for example, scene changes using very low-level cues and to extract key frames at those instances to form a storyboard for easy video access. Some institutions are running projects to detect subtitles in the video scene and create a textual annotation. This task is very difficult, given a sequence of real scenes with moving backgrounds and so on. Even more ambitious tasks such as finding real patterns in real movies (tracing the course of the ball in a soccer match, for example) are still far from being achieved.[2]

There are also some starting multimedia projects within the IST program which at this stage can not be compared to MUMIS, since we still lack first intermediate results.

3. Domain-specific Annotations

The consortium of the project has chosen soccer as the domain to test and apply the algorithms to be developed. There are a number of reasons for this choice: availability of people willing to help in analyzing user requirements, existence of many information sources in several languages[3], and great economic and public interest.

The principles and methods derived from this domain can be applied to other as well. This has been shown already in the context of text-based Information Extraction (IE), for which methodologies for a fast adaptivity to new domains have been developed (see the MUC conferences and (Neumann et al., 2000)). And generally speaking the use of IE for automatic annotation of multimedia document has the advantage of providing, besides the results of the (shallow) syntactic processing, accurate semantic (or content/conceptual) information (and thus potential annotation) for specific predefined domains, since a mapping from the linguistically analyzed relevant text parts can be mapped onto an unambiguous conceptual description. This topic has already been object of a workshop discussing the relations between IE and Corpus Linguistics (McNaught, 2000). Thus in a sense it can be assumed that IE is supporting the word sense disambiguation task.

[2] The URLs of the projects mentionned above are given in the bibliography at the end of this paper.

[3] We would like to thank at this place the various institutions making available various textual, audio and video data.

It is also commonly assumed (see among others (Cunningham, 1999)) that IE occupies an intermediate place between Information Retrieval (with few linguistic knowledge involved) and Text Understanding (involving the full deep linguistic analysis and being still not realized for the time being.). IE being robust but offering only a partial (but mostly accurate) syntactic and content analysis, it can be said that this language technology is actually filling the gap between available low-level annotated/indexed documents and corpora and the desirable full content annotation of those documents and corpora. This is the reason why MUMIS has chosen this technology for providing automatic annotation (at distinct linguistic and domain-specific levels) of multimedia material.

4. Processing Steps and Annotation Levels in MUMIS

4.1. Collection of Data

The MUMIS project is about automatic indexing of videos of soccer matches with formal annotations and querying that information to get immediate access to interesting video fragments. For this purpose the project chose the European Football Championships 2000 in Belgium and the Netherlands as its main database. A major project goal is to merge the formal annotations extracted from textual and audio material (including the audio part of videos) on the EURO 2000 in three languages: English, German, Dutch. Therefore, the project has collected raw textual and audio material of different sort from different countries as well as the corresponding videos for building a representative corpus for training and testing the automatic indexing of videos:

1. Reports from Newspapers (reports about specific games, general reports) which is classified as free texts (FrT)
2. Tickers, close captions, Action-Databases which are classified as semi-formal texts (SFT)
3. Formal descriptions about specific games which are classified as formal texts (FoT)
4. Audio material recorded from radio broadcasts
5. Video material recorded from TV broadcasts

The texts and audio information have been gathered for the three indexing languages. Textual material will also be gathered for Swedish, since this language will be used as a neutral control language in the user interface of the system. All the data are available as XML-structured documents (indicating their language, source and type).

Since the information contained in formal texts can be considered as a database of true facts, they play an important role within MUMIS. But nevertheless they contain only few information about a game: the goals, the substitutions and some other few events (penalties, yellow and red cards). So there are only few time points available for indexing videos. Semi-formal texts (SFT), like live tickers on the web, are offering much more time points sequences, related with a higher diversity of events (goals scenes, fouls etc,) and seem to offer the best textual source for our purposes. Nevertheless the quality of the texts of online tickers is often quite poor. Free texts, like newspapers articles, have a high quality but the extraction of time points and their associated events in text is more difficult. Those texts also offer more background information which might be interesting for the users (age of the players, the clubs they are normally playing for, etc.). Figures 1 and 2 in section 9. show examples of (German) formal and semi-formal texts on one and the same game.

4.2. Media Pre-Processing

Media material has been delivered in various formats (AudioDAT, AudioCassettes, Hi-8 video cassettes, DV video cassettes etc) and qualities. All audio signals (also those which are part of the video recordings) are digitized and stored in an audio archive. Audio digitization is done with 20 kHz sample frequency, the format generated is according to the de-facto wav standard. For digitization any available tool can be used such as SoundForge.

Video information (including the audio component) of selected games have been digitized into MPEG1 streams first. Later it will be encoded in MPEG2 streams. While the quality of MPEG1 is certainly not satisfying to the end-user, its bandwidth and CPU requirements are moderate for current computer and network technology. The mean bit rate for MPEG1 streams is about 1.5 Mbps. Current state-of-the-art computers can render MPEG1 streams in real time and many network connections (Intranet and even Internet) can support MPEG1. MPEG2 is specified for about 3 to 5 Mbps. Currently the top-end personal computers can render

MPEG2, but MPEG2 is not yet supported for the most relevant player APIs such as JavaMediaFramework or Quicktime. When this support is given the MUMIS project will also offer MPEG2 quality.

For all separate audio recordings as for example from radio stations it has to be checked whether the time base is synchronous to that one of the corresponding video recordings. In case of larger deviations a time base correction factor has to be estimated and stored for later use. Given that the annotations cannot be created with too high accuracy a certain time base deviation will be accepted. For part of the audio signals manual transcriptions have to be generated to train the speech recognizers. These transcripts will be delivered in XML-structured files.

Since keyframes will be needed in the user interface, the MUMIS project will develop software that easily can generate such keyframes around a set of pre-defined time marks. Time marks will be the result of information extraction processes, since the corresponding formal annotations is referring to to specific moments in time (see examples of time marks extracted from formal texts in figure 4 in section 9.). The software to be written has to extract the set of time marks from the XML-structured formal annotation file and extract a set of keyframes from the MPEG streams around those time marks. A set of keyframes will be extracted around the indicated moments in time, since the estimated times will not be exact and since the video scenes at such decisive moments are changing rapidly. There is a chance to miss the interesting scene by using keyframes and just see for example spectators. Taking a number of keyframes increases the chance to grab meaningful frames.

4.3. Multilingual Automatic Speech Recognition

Domain specific language models will be trained. The training can be bootstrapped from written reports of soccer matches, but substantial amounts of transcribed recordings of commentaries on matches are also required. Novel techniques will be developed to interpolate the base-line language models of the Automatic Speech Recognition (ASR) systems and the domain specific models. Moreover, techniques must be developed to adapt the vocabularies and the language models to reflect the specific conditions of a match (e.g., the names players have to be added to the vocabulary, with the proper bias in the language model). In addition, the acoustic models must be adapted to cope with the background noise present in most recordings.

Automatic speech recognition of the sound tracks of television and (especially) radio programmes will make use of closed caption subtitle texts and information extracted from formal texts to help in finding interesting sequences and automatically transcribing them. Further, the domain lexicons will help with keyword and topic spotting. Around such text islands ASR will be used to transcribe the spoken soundtrack. The ASR system will then be enriched with lexica containing more keywords, to increase the number of sequence types that can be identified and automatically transcribed.

4.4. Multilingual Domain Lexicon Building

All the collected textual data for the soccer domain are used for building the multilingual domain lexicons. This data can be in XML, HTML, plain text format, etc. A number of automatic processes are used for the lexicon building, first on a monolingual and secondly on a multilingual level. Manual browsing and editing is taking place, mainly in order to provide the semantic links to the terms, but also for the fine-tuning of the lexicon according to the domain knowledge.

Domain lexicons are built for four languages, namely English, German, Dutch and Swedish. The lexicons will be delivered in a fully structured, XML-compliant, TMX-format (Translation Memory eXchange format). For more information about the TMX format see http://www.lisa.org/tmx/tmx.htm.

We will also investigate how far EUROWORDNET resources (see http://www.hum.uva.nl/ ewn/) can be of use for the organization of the domain-specific terminology.

4.5. Building of Domain Ontology and Formal Annotations

The project is currently building an ontology for the soccer domain, taking into consideration the requirements of the information extraction and merging components, as well as users requirements. The ontology will be delivered in an XML format[4].

[4]There are still on-going discussions within the project consortium wrt the best possible encoding format for the domain ontology, the alternative being reduced probably to RDFS, OIL and IFF, see respec-

In parallel to building the ontology an event table is being described. It contains the major event types that can occur in soccer games and their attributes. This content of the table is matching with the content of the ontology. The event table is a flat structure and guides the information extraction processes to generate the formal event annotations. The formal event annotations build the basis for answering user queries. The event table is specified as an XML schema to constrain the possibilities of annotation to what has been agreed within the project consortium. The time marks presented in figure 4 in section 9. are also giving examples of such events.

4.6. Generation of Formal Annotations

The formal annotations are reflecting the typical output of the IE systems being used in MUMIS, i.e.instantiated domain-specific templates. The slots to be filled by the systems are basically entities (player, teams etc.), relations (player_of, opponents etc.) and events (goal, substitution etc.), which are all derived from the current version of the domain ontology. Figure 3 in section 9. is giving an example of an uninstantiated `team' (i.e. entity) template. All the templates associated with an event are including a time slot to be filled if the correpsonding information is available in a least one of the sources consulted during the IE procedure.

The IE systems include a lookup procedure in the domain lexicon (the corresponding monolingual part of it). The systems will have to apply to distinct sources (FoT, FrT etc.) but they are concerned with achieving consistency in the IE result on distinct sources about the same event (game): this is the task of the merging tools, described below. The results of each IE system are mapped onto XML.

At the end of the process, only those formal annotation will be delivered which are relevant for the purpose of querying the videos, but one has to be aware of that all the other information used and provided by the IE systems is also at some place available and can be used for annotating as well textual documents with linguistic and domain-specific information[5].

4.7. The Merging Tool

When distinct formal annotations are available on one event (being due to the variety of IE-processed sources or the variety of languages, or a combination of both), a merging of those is necessary, in order to avoid both inconsistencies and redundancies in the annotations.

The merging tools used in MUMIS will not only apply to the results of IE, but will also take into consideration some general representation of the domain-knowledge in order to filter out some annotations generated in the former phases.

The use of general representations[6] (like domain frames) might also support a better sequential organization of some event templates in larger scenarios. It will also allow to induce some events which are not explicitly mentioned in the sources under consideration (or which the IE systems might not have detected).

The merged formal annotations will be mapped onto XML.

4.8. User Interface Building

The user first will interact with a web-portal to start a MUMIS query session. An applet will be down-line loaded in case of showing the MUMIS demonstration. This applet mainly offers a query interface. The user then will enter a query that either refers to metadata, formal annotations, or both. The MUMIS on-line system will search for all formal annotations that meet the criteria of the query. In doing so it will find the appropriate meta-information and/or moments in some media recording. In case of meta-information it will simply offer the information in scrollable text widgets. This will be done in a structured way such that different type of information can easily be detected by the user. In case that scenes of games are the result of queries about formal annotations the user interface will first present selected video keyframes as thumbnails with a direct indication of the corresponding metadata.

The user can then ask for more metadata about the corresponding game or for more media data. It has still to be decided within the project whether several layers of media data zooming in and out are useful to satisfy the user or whether the step directly

---

tively, and among others, http://www.w3.org/TR/rdf-schema/, http://www.oasis-open.org/cover/oil.html and http://www.ontologos.org/IFF/The%20IFF%20Language.html

[5]This issue has been discussed in more details in (Declerck and Neumann, 2000).

[6]Like for example the Type Description Language (TDL), a formalism supporting all kind of operations on (typed) features as well as multiple inheritance, see (Krieger and Schaefer, 1994).

to the corresponding video fragment is offered. All can be invoked by simple user interactions such as clicking on the presented screen object. Playing the media means playing the video and corresponding audio fragment in streaming mode requested from a media server.

The user interface will be offered in four localized versions (English, Dutch, German, Swedish). Swedish is a test query language in order to demonstrate the multilingual user interface.

5. Standards for Multimedia Content

MUMIS is looking for a compliance with existing standards in the context of the processing of multimedia content on the computer. So all the data to be interchanged within the context of the project (internaaly or externally) will be as far as possible encoded in XML (DTD and Schemas). And MUMIS will adhere to emerging standards such as MPEG4, which defines how different media objects will be decoded and integrated at the receiving station, and MPEG7, which is about defining standards for annotations which can be seen as multimedia objects. Further, MUMIS will also maintain awareness of international discussions and developments in the aerea of multimedia streaming (RTP, RTSP, JMF...), and will follow the discussions within the W3C consortium and the EBU which are also about standardizing descriptions of media content.

In the course of the project MUMIS will discuss as well the compliance to annotation frameworks, as proposed for example by Bird and Liberman or those presented at the MMA Workshop (GDA, UNL, ISLE).

6. Role of MUMIS for the Annotation of Multimedia Content

To conclude, we would like to list the points where we think MUMIS can, directly or indirectly, contribute to the goals of the First Multimedia Annotation Workshop. In the description of the project, we have shown that MUMIS is an innovative multimedia indexing project which:

- uses multimedia (MM) and multilingual information sources;
- carries out multimedia indexing by applying information extraction to a well-delineated domain and using already existing information as constraints;
- uses and extends advanced language technology to automatically create formal annotations for MM content;
- merges information from many sources to improve the quality of the annotation database;
- application of Automatic Speech Recognition (ASR) to emotionally-coloured spontaneous speech in a special environment and under constraints given by earlier extracted information;
- application of IE to the output of ASR and the combination of this with already existing knowledge;
- definition of a complex information annotation structure, which is stored in a standard document type definition (DTD);
- integration of new methods into a query interface which is guided by domain knowledge (ontology and multilingual lexica).

So in a sense MUMIS is contributing in defining semantic structures of multimedia contents, at the level proposed by domain-specific IE analysis. The full machinery of IE, combined with ASR (and in the future with Image Analysis) can be used for multimedia contents development and so efficiently support cross-media (and cross-lingual) information retrieval and effective navigation within multimedia information interfaces.

7. bibliography

8. References

Appelt, Doug E., 1999. An introduction to information extraction. *AI Communications*, 12.

Bird, Steven and Mark Liberman, 2001. A formal framework for linguistic annotation. *Speech Communication*.

Bontcheva, K., H. Brugman, A. Russel, P. Wittenburg, and H. Cunningham, 2000. An Experiment in Unifying Audio-Visual and Textual Infrastructures for Language Processing R&D. In *Proceedings of the Workshop on Using Toolsets and Architectures To Build NLP Systems at COLING-2000*. Luxembourg. `http://gate.ac.uk/`.

Broeder, Daan, Hamish Cunningham, Nancy Ide, David Roy, Henry Thompson, and Peter Wittenburg (eds.), 2000. *Meta-Descriptions and Annotation Schemes for Multimodal/Multimedia Langauge Resources LREC-2000*.

Brugman, H., K. Bontcheva, P. Wittenburg, and H. Cunningham, 1999. Integrating Multimedia and Textual Software Architectures for Language Technology. Technical report mpi-tg-99-1, Max-Planck Institute for Psycholinguistics, Nijmegen, Nethedlands.

Cunningham, Hamish, 1999. An introduction to information extraction. *Research memo CS - 99 - 07*.

Declerck, Thierry and G. Neumann, 2000. Using a parameterisable and domain-adaptive information extraction system for annotating large-scale corpora? In *Proceedings of the Workshop Information Extraction meets Corpus Linguistics, LREC-2000*.

Humphreys, Kevin, R. Gaizauskas, S. Azzam, C. Huyck, B. Mitchell, H. Cunningham, and Y. Wilks, 1998. University of sheffield: Description of the lasie-ii system as used for muc-7. In SAIC (ed.), *Proceedings of the 7th Message Understanding Conference, MUC-7*. http://www.muc.saic.com/: SAIC Information Extraction.

Kennedy, Christopher and B. Boguraev, 1996. Anaphora for everyone: Pronominal anaphora resolution without a parser. In *Proceedings of the 16th International Conference on Computational Linguistics, COLING-96*. Copenhagen.

Krieger, Hans-Ulrich and U. Schaefer, 1994. $\mathcal{TDL}$—a type description language for constraint-based grammars. In *Proceedings of the 15th International Conference on Computational Linguistics, COLING-94*.

Lappin, Shalom and H-H. Shih, 1996. A generalized algorithm for ellipsis resolution. In *Proceedings of the 16th International Conference on Computational Linguistics, COLING-96*. Copenhagen.

McNaught, John (ed.), 2000. *Information Extraction meets Corpus Linguistics, LREC-2000*.

Mitkov, Ruslan, 1998. Robust pronoun resolution with limited knowledge. In *Proceedings of the 17th International Conference on Computational Linguistics, COLING-98*. Montreal.

MUC (ed.), 1995. *Sixth Message Understanding Conference (MUC-6)*. Morgan Kaufmann.

MUC (ed.), 1998. *Seventh Message Understanding Conference (MUC-7)*. http://www.muc.saic.com/: SAIC Information Extraction.

Neumann, Guenter, R. Backofen, J. Baur, M. Becker, and C. Braun, 1997. An information extraction core system for real world german text processing. In *Proceedings of the 5th Conference on Applied Natural Language Processing, ANLP-97*.

Neumann, Guenter, C. Braun, and J. Piskorski, 2000. A divide-and-conquer strategy for shallow parsing of german free texts. In *Proceedings of the 6th Conference on Applied Natural Language Processing, ANLP-00*.

Piskorski, Jakub and G. Neumann, 2000. An intelligent text extraction and navigation system. In *Proceedings of the 6th Conference on Recherche d'Information Assistée par Ordinateur, RIAO-2000*.

Project URLs:

```
COL: http://www.cs.columbia.edu/ hjing/
      sumDemo
EUR: http://www.foyer.de/euromedia/
GDA: http://www.csl.sony.co.jp/person/
      nagao/gda/
INF: http://www.informedia.cs.cmu.edu/
ISI: http://www.wins.uva.nl/research/
      isis/isisNS.html
ISLE: http://www.ilc.pi.cnr.it/EAGLES/
      ISLE_Home_Page.htm
NSF: http://www.nsf.gov./od/lpa/news/press/
      pr9714.htm
OLI: http://twentyone.tpd.tno.nl/olive
POP: http://twentyone.tpd.tno.nl/popeye
SUR: http://www-rocq.inria.fr/ nastar/MM98/
      node1.html
THI: http://www.dcs.shef.ac.uk/research/
      groups/spandh/projects/thisl
UMA: http://ciir.cs.umass.edu/research/
UNL: http://www.ias.unu.edu/research_prog/
      science_technology/
      universalnetwork_language.html
VIR: http://www.virage.com/
```

9. Annex

England - Deutschland 1:0 (0:0)
England: Seaman (2,5) - G. Neville (3,5), Keown (3), Campbell (2), P. Neville (4,5) - Ince (3,5), Wise (5) - Beckham (4), Scholes (3) - Shearer (3), Owen (5) - Trainer: Keegan
Deutschland: Kahn (2) - Matthaeus (3) - Babbel (3,5), Nowotny (2,5) - Deisler (3), Hamann (2,5), Jeremies (3,5), Ziege (3,5) - Scholl (5) - Jancker (4), Kirsten (5) - Trainer: Ribbeck
Eingewechselt: 61. Gerrard fuer Owen, 72. Barmby fuer Scholes - 70. Rink fuer Kirsten, 72. Ballack fuer Deisler, 78. Bode fuer Jeremies
Tore: 1:0 Shearer (53., Kopfball, Vorarbeit Beckham)
Schiedsrichter: Collina, Pierluigi (Viareggio), Note 2 - bis auf eine falsche Abseits-Entscheidung souveraen und sicher
Zuschauer: 30000 (ausverkauft)
Gelbe Karten: Beckham - Babbel, Jeremies

Figure 1: Example of a so-called formal text, where one can see that only 5 distinct time points can be extracted, concerning the player subsitutions ("Eingewechselt") and one goal ("Tore").

Gruppe A: England - Deutschland 1:0 (0:0)
7. Ein Freistoss von Christian Ziege aus 25 Metern geht ueber das Tor.
12. Ziege flankt per Freistoss in den Strafraum und Jeremies versucht es per Kofball, verfehlt den Kasten jedoch deutlich.
16. Scholes flankt gefaehrlich von der Torauslinie in den Fuenfmeterraum, doch Ziege hat aufgepasst und kann klaeren.
18. Hamann versucht es mit einem Distanzschuss aus 20 Metern, aber Seaman ist auf dem Posten.
23. Scholl mit einer Riesenchance: Nach Zuspiel von Hamann rennt er in den englischen Strafraum, wird jedoch gleich von drei Seiten bedraengt und kommt nur zu einem unplazierten Schuss, den Seaman sicher abfangen kann.
27. Jancker spielt auf Ziege, dessen Schuss von der Strafraumgrenze kann von Seaman abgefangen werden.
35. Michael Owen kommt nach Flanke von Philip Neville voellig frei vor dem deutschen Tor zum Kopfball, doch Kahn kann zum ersten Mal sein Koennen unter Beweis stellen und rettet auf der Linie.
43. Kahn zum zweiten: Beckham flankt auf Scholes, der zieht ab in den rechten Winkel, aber der deutsche Keeper verhindert erneut die englische Fuehrung.
47. Christian Zieges Freistoss aus 20 Metern geht einen halben Meter ueber das Tor.
53. Beckham flankt per Freistoss an der deutschen Abwehr vorbei auf den Kopf von Alan Shearer, der voellig freistehend zum 1:0 fuer die Englaender verwandelt.
58. Scholl wird von Matthaeus bedient, aber sein Schuss geht aus halbrechter Position um Zentimeter am Tor vorbei.
65. Seaman kann nach einem Eckball vor Kirsten klaeren, der Nachschuss von Jancker geht knapp am Tor vorbei. Riesenmoeglichkeit fuer die DFB-Elf.

Figure 2: Example of a so-called semi-formal text, where one can see that here more time points are available, and that those can be complementary to the time points to be extracted from formal texts. So, already at this level, a unification or merging of extracted time points is necessary.

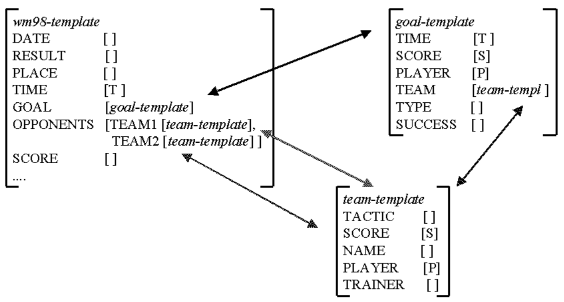

Figure 3: An 'entity' template for the soccer domain: the TEAM-template and its embedding in various level of template definition (*event* and *scenario* templates). As can be seen, information-sharing between the templates is here supported and represented by variables.

<DOM=SOCCER TYPE=SUBSTITUTION TIME=61:00 SYN=NP PLAYER_OUT=Owen PLAYER_IN=Gerrard>

<DOM=SOCCER TYPE=GOAL GOAL_TYPE=Kopfball TIME=53:00 SCORE=1:0 SYN=NP PLAYER=Shearer PREPARATION=Beckham>

Figure 4: Two examples of time-marks and their associated events (substitution and goal) automatically extracted from the formal text above.