University of Padova
Department of Electronics and Computer Science
Via Gradenigo, 6/a
35131 Padova
Italy

{melo,orio}@dei.unipd.it

The use of melodic segmentation and normalization for content-based retrieval of musical data

Massimo Melucci and Nicola Orio

Information Management Systems research group

Computer Music research group

Karlsruhe, 4 February 2000

Presentation Outline

- **Motivation:** why *musical* information retrieval?
- **Approach:** applying information retrieval to music
- **Segmentation:** finding “lexical units” in musical language
- **Normalization:** stemming of musical phrases
- **Experimental study:** evaluation of performances
- **Hypermusic:** browsing musical documents
- **SMILE:** a **S**ystem for **M**usical **I**nformation retrieva**L** **E**nvironments

Motivation 1/3

- Availability of musical material on the web is growing. The diffusion of old (e.g. MIDI) and new (e.g. MP3) formats, and the development of systems with multimedia capabilities gave a boost to the diffusion of music on the Internet.

- Potential users may have a different musical knowledge and background as well as a different expertise. This implies the need of a novel user-friendly system allowing the musical information retrieval to all the users categories (from the music scholar to the untrained pure listener).

Motivation 2/3

- Catalogue-based access requires a good knowledge of the musical literature. Hence it can be very effective only for expert users.
- String-matching algorithms are sometimes inefficient and ineffective (e.g. **`grep`** in Unix). Musical language uses many *variants* of the same material, which cannot be taken into account by simple string-matching techniques.
- Musical information retrieval should be content-based. Yet it is still unclear which kind of content, if any, musical works do convey.

Motivation 3/3

- Investigation of use of standard, efficient and effective information retrieval techniques for musical information retrieval. Information retrieval has dealt with the representation and the disclosure of content from its early days.

- Organization, integration, and access to the digitized version of music documents are necessary within *digital libraries*.

- At the Department of Electronics and Computer Science a feasibility study was conducted for the ADMV (*Archivio Digitale della Musica Veneta*, Digital Library for Venetian Music) project.

Approach 1/5

- Combination of Information Retrieval (IR) and Computer Music techniques. Instead of defining ad-hoc algorithms, we integrate music and sound analysis with IR well-known techniques.

- We speak about "indexing and retrieval of musical documents matching musical queries".

- Use of standard IR techniques requires the identification of lexical units like words in textual documents. Provided that "the concept of a lexical unit corresponding to a word may have more meaning in some musical works than others" (McLane, ARIST, 1996).

Approach 2/5

- A duality is supposed between textual and musical languages, where "musical phrases" play the same role of textual words, that is they are *lexical units*.

- Music does not have explicit separators, which text has (i.e. blanks, commas, dots). Hence, it is necessary to perform an automatic *segmentation* of musical documents and queries for highlighting their lexical units.

- Musical phrases may not be equal, but they may be perceived similar by the listener, that is they may carry the same content. Hence, it is necessary to perform an automatic *normalization* of the musical phrases for taking into account the similarities among different phrases.

Approach 3/5

Musical Indexing vs. Textual Indexing

Musical Indexing	Textual Indexing
Segmentation	Words Recognition
Normalization	Stemming
Index Creation	Index Creation
Hypermusic	Hypertext

Segmentation and normalization algorithms extract content-descriptive musical phrases from both documents and queries. The information retrieval techniques are employed to index, retrieve, and rank by score the musical documents judged relevant to the final users' information need.

Approach 4/5

- Two types of retrieval functions are provided – a *querying function* with an interface allowing the user to play the query, and a *browsing function* to navigate an automatically constructed hypermusic which is made of documents and phrases.

- The term "hypermusic" indicates an hypermedia of which nodes are pieces of music.

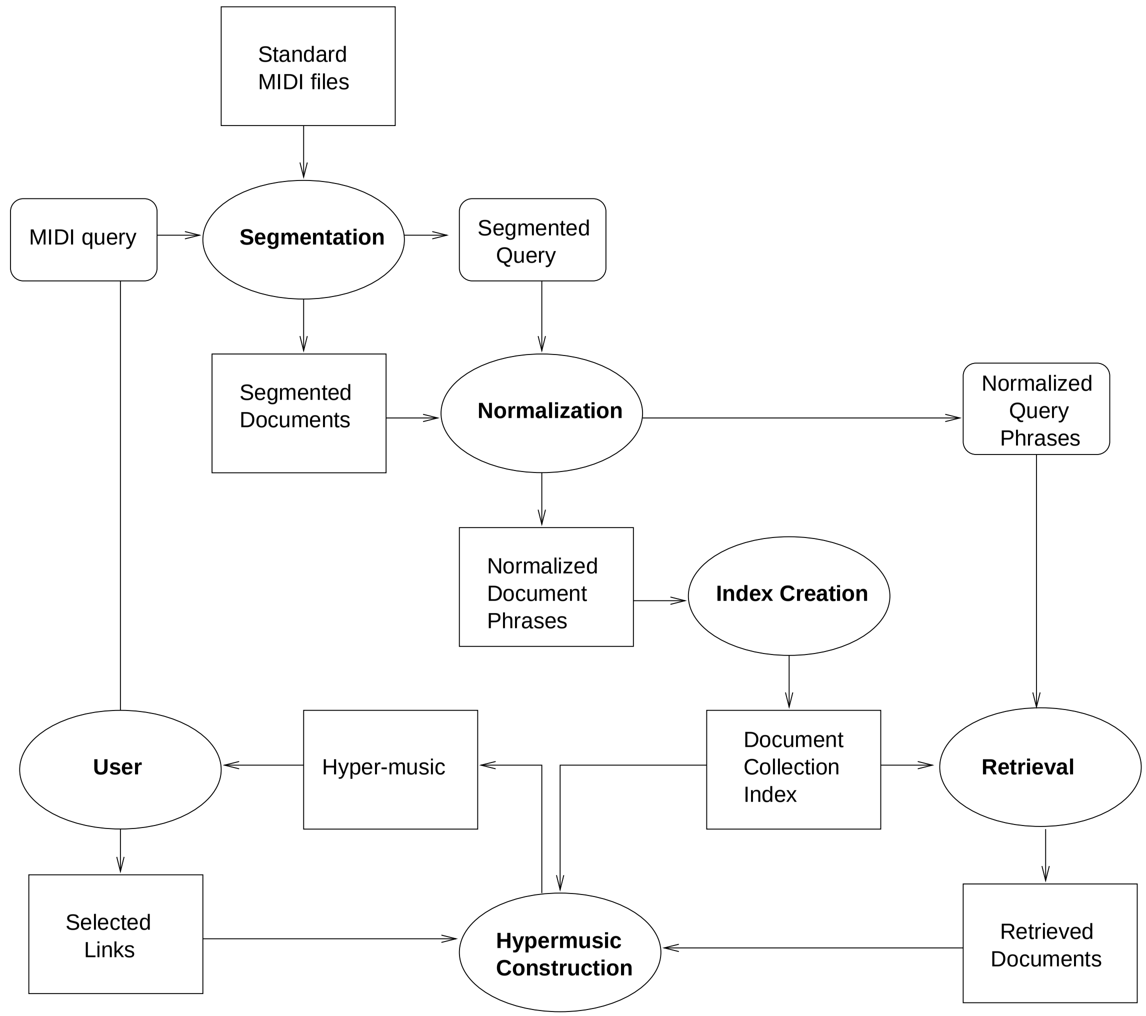
- Hypermusic links are constructed automatically and are based on content, i.e. the algorithms used to detect links employ the index-based description of the musical document content.

Approach 5/5

Documents and queries, originally in MIDI format, are segmented and normalized.

Retrieval
is performed using standard IR matching algorithms between phrases.

The user is then able to navigate in a network of musical documents and musical phrases.



Segmentation 1/4

- Musical notation, as well as musical performances, lack of explicit separators. Musical *rests* have a different role than *blanks* in text.

- Nevertheless, listeners perceive the presence of lexical units in the musical language. Moreover, musicians' expressiveness is strictly related to these lexical units.

Musical language offers a number of cues to performers and listeners, helping them in highlighting the lexical units.

Music theorists (e.g. Narmour; Lerdhal and Jackendoff) proposed some models that correlates these lexical units with the musical structure, the rhythm, the harmony. and the melody.

Segmentation 2/4

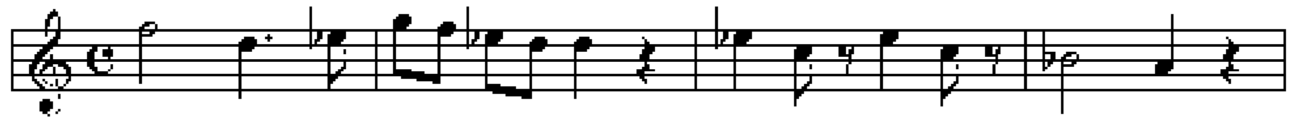
With the term "lexical unit", we can refer to different characteristics of the musical language, that is a *rhythmic cell*, or a *cadenza*, or a *melodic excerpt*.

In our work we took into account two key considerations:

- Segmentation in lexical units has to performed both on documents and on queries.
- A typical end-user may not have a good knowledge of the music language, and he may be able to formulate only a simple query.

Hence we chose to consider only the informations regarding the melody; and that's why we call the lexical units *musical phrases*.

Segmentation 3/4



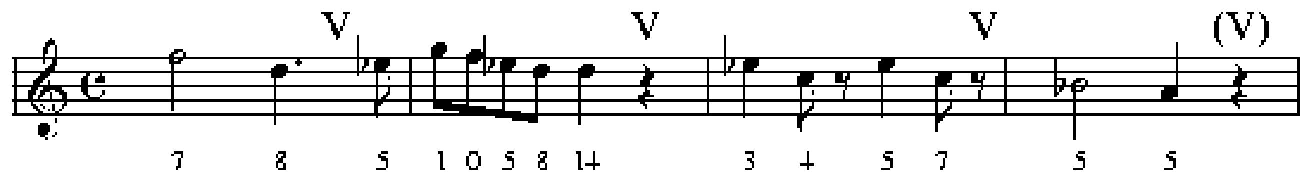
Searching boundaries in the melody through weighting.

1. musical intervals – If intervals are different (e.g. a fifth versus a third), then a given value is added to the two weights, depending on the relationship between intervals (bigger interval = higher weight)

2. note durations – If durations are different a given value is added to both intervals; the first note will always have an higher weight, weights depend on the relationship between durations

3. musical rests – In this case a value is added to the weight of the note preceding the rest

Segmentation 4/4



Each note has a given probability to be the last one of a musical phrase. The probability is supposed to be proportional to the weight given to each note.

The detection of boundaries in the melodic surface is performed by analyzing the presence of maxima in the weight function.

We highlighted two different lexical units:

1. Phrases are melodic surfaces bounded by two subsequent maxima

2. Periods are groups of subsequent phrases bounded by a maximum among maxima

We used periods just in the testing phase of the system.

Normalization 1/5

- Two phrases with the same pitch contour may be perceived similar, even if they are played with a different Tempo.
- Two phrases with the same rhythm may be perceived similar, even if they have different pitches.
- Two phrases are perceived similar when the second is a modulation of the first one.
- Grace notes change the melodic profile, which is still perceived similar to the one without grace notes.

Normalization 2/5

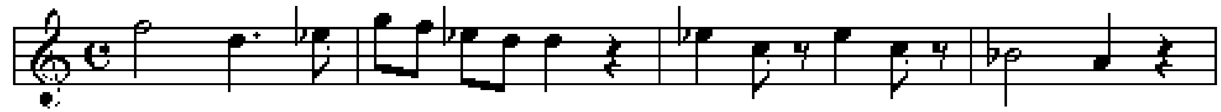
We tried to overcome partially these problems by performing a normalizations of the pitches and durations of the musical phrases.

Different normalizations were combined to obtain an increasing generality of the musical phrases. Hence it is more probable that two different phrases perceived as similar will become equal after the normalization.

The normalization is the analogous of the *stemming* in textual information retrieval.

Note: it is likely to go further with the duality and to find the analogous of *stop words*, but interviewed musicians and theorists could not highlight something like *stop musical phrases*.

Normalization 3/5



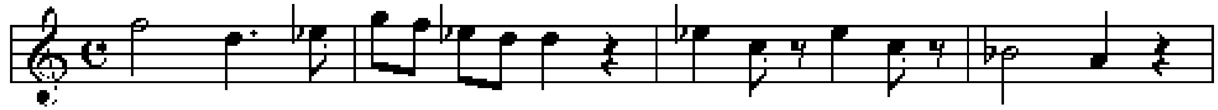
PT: Pitch Transposition. All the phrases are transposed in order to begin with a note of pitch C5.

TDN: Pitch Transposition and Duration Normalization. PT + note durations are expressed in multiples of their Greatest Common Divisor. Relationships among durations are important.

NDN: Pitch Normalization and Duration Normalization. DN + pitches are quantized in a number of different size of musical intervals. The trend of the melodic surfaces becomes significant.

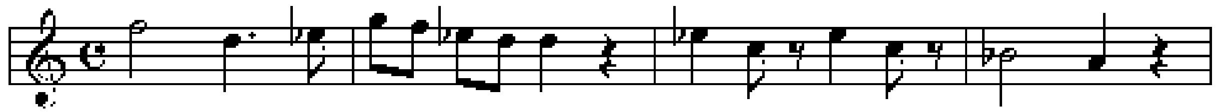
NDR: Pitch Normalization and Duration Removal. PNDN + no information about duration is maintained. To test the importance of timing versus interval sequence.

Normalization 4/5



Visual representation of the score after applying the four different kinds of normalization.

Normalization 5/5



Internal representation of the score after applying the four different kinds of normalization.

Normalization	Phrases
PT	C524A418 C56E56D56C56B46B424 C512A412C512A412 C524B412
PTDN	C54A43 C51E51D51C51B41B44 C51A41C51A41 C52B41
PNDN	O4N3 O1P1N1N1N1O4 O1N1P1N1 O2N1
PNDR	ON OPNNNO ONPN ON

Experimental study: Test Data.

- Documents: 419 pieces in standard MIDI files about 7 authors (Bach, Beethoven, Chopin, Händel, Listz, Mozart, Tchaikovskij).
- Queries:
  - Type A: 419 incipit of 1, 2, 3, 4, 5, 10 tokens for each normalization
  - Type B: 15 short melodies played by a musician "by hand" and then normalized
- Indexing methods: only phrases and only periods.
- Retrieval model: Vector-space model.

Experimental study: Quantitative Analysis

- Based on Type A queries (incipit).
- Two variables are observed: the number of retrieved documents (NDOCS), and the number of different retrieved authors (NAUTHORS).
- Two tests: using phrases only, using periods only.
- As the normalization becomes higher or query length increases (i.e. from PT to PNDR), NDOCS and NAUTHORS increases.
- With periods NDOCS and NAUTHORS are less than with phrases.

Experimental study: Qualitative Analysis

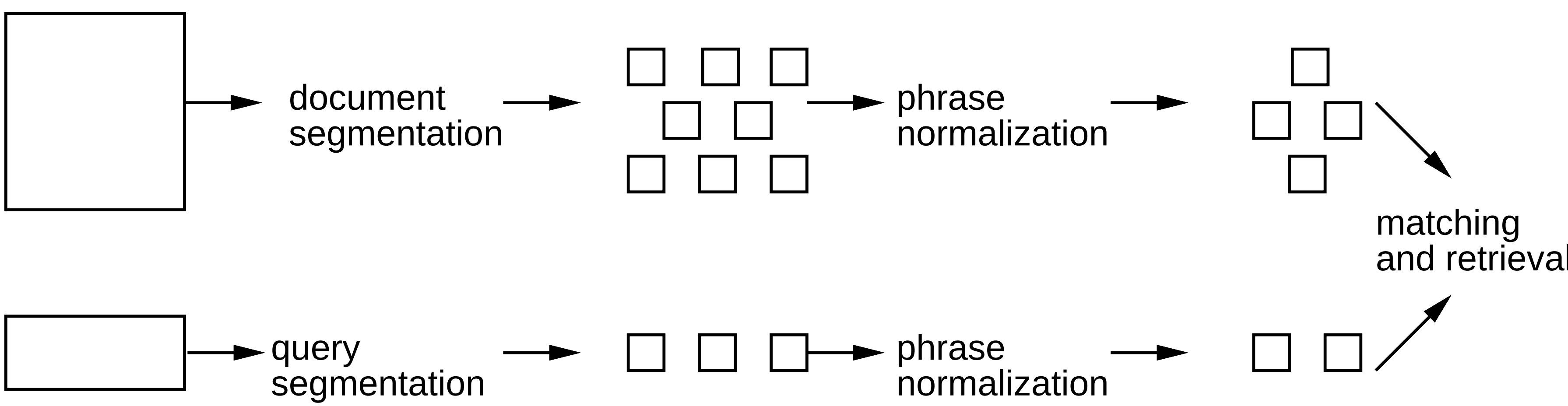
- Based on Type B queries ("by hand").

- Queries using pieces that were present in the collection always gave that same piece with the highest RSV, independently on normalization. This is particularly important if document finding is performed instead of document retrieval.

- Most of the retrieved melodies shared the same *temper* of the queries, even if this temper is related to the musical style of the different composers.

Hypermusic 1/6

Insofar, the methodology of the musical information retrieval system can be represented as follows:



We propose a novel approach for querying a musical digital library, based on the concept of *similarity* among musical documents and musical phrases.

Hypermusic 2/6

The proposed approach allows the end-user to navigate inside an hypermedia, which nodes are pieces of music, and links among hypertexts are automatically created between:

- Documents (DD links): two documents are judged similar if they share the same representative phrases.
- Phrases (TT links): musical phrases are judged similar if they are representative of the same documents.
- Documents and phrases (DT links): to phrases which are most representative of the document.
- Phrases and documents (TD links): to documents for which the phrase is representative.

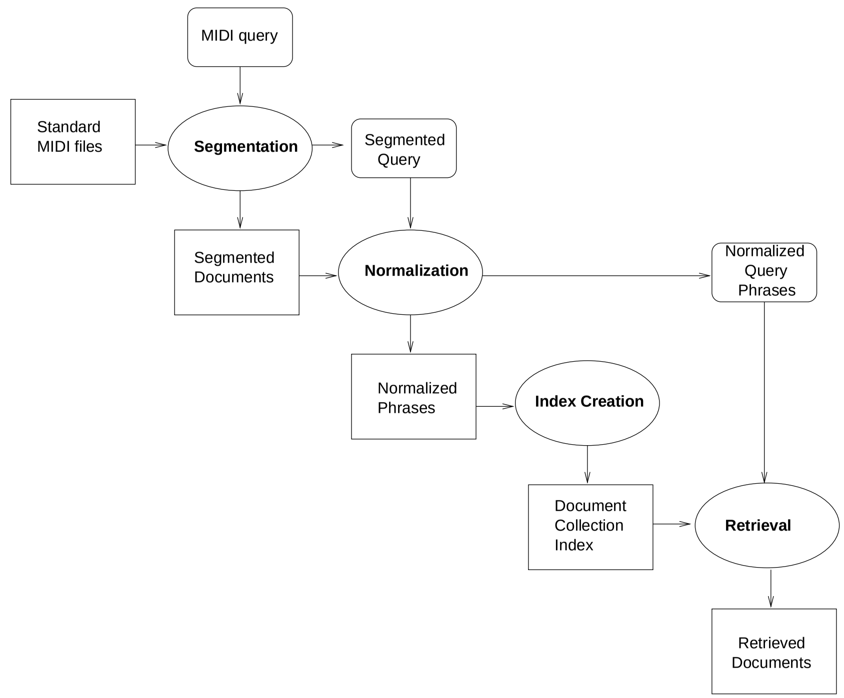
Hypermusic 3/6

a: The starting point is a collection of musical documents. Documents are at the D level

b: Next step is automatic segmentation and normalization for extracting musical phrases. Phrases are at the T level.

c: Indexing produces an index of musical phrases being associated to the musical documents.

d: Automatic link construction involves four types of link and differs whether link is among nodes of the same level, i.e. DD or TT links, or among nodes of different levels, i.e. TD or DT links.



Hypermusic 4/6

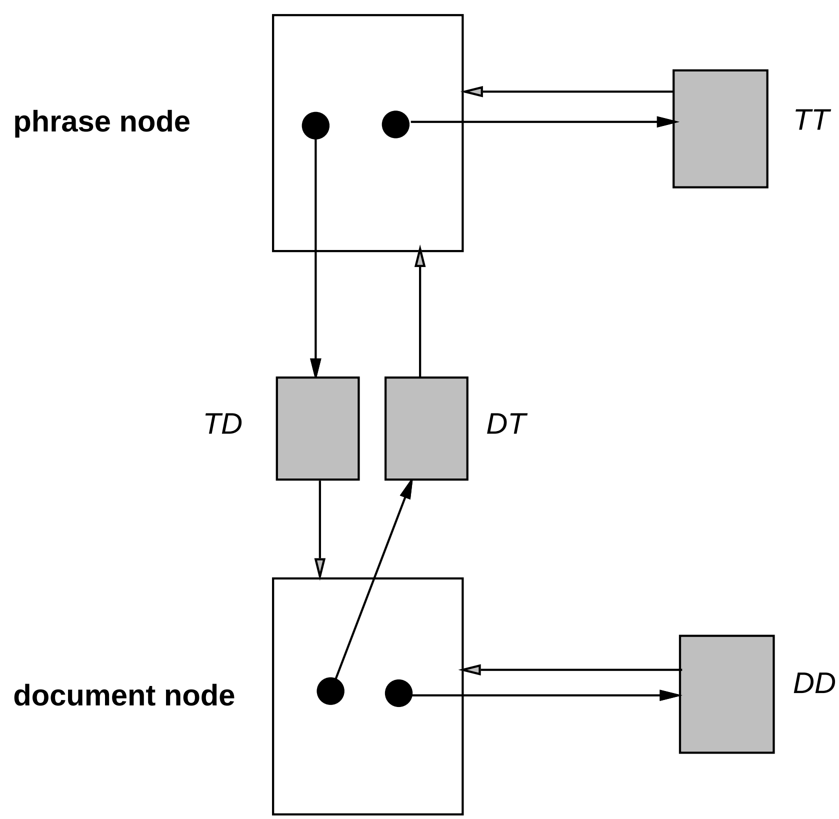
After having retrieved a number of documents, the user can start navigating inside hypermusic, following:

DD links: to retrieve documents being similar to those retrieved, but judged not relevant by the system and hence unretrieved.

DT links: to look up phrases describing the document content, especially the phrases being absent within the starting query.

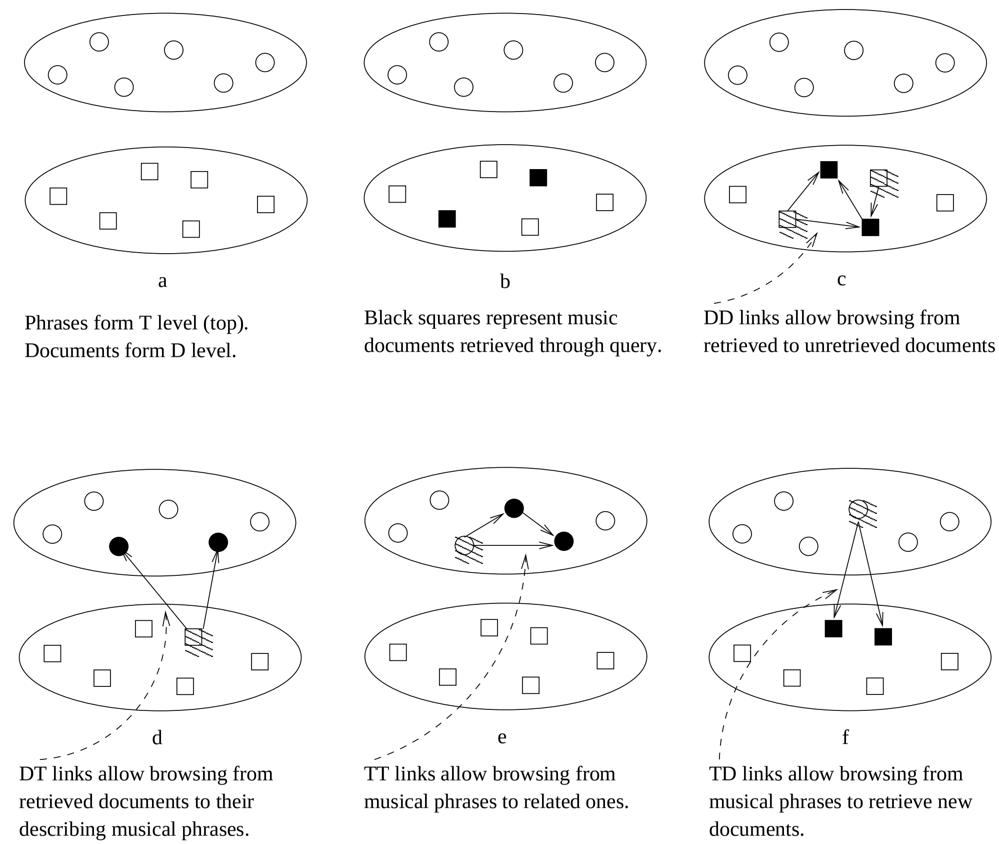
TT links: to reformulate the query whenever the user wants to use similar phrases previously ignored, for retrieving new documents.

TD links: to retrieve the documents indexed by a phrase without formulating an ad-hoc query – moving along a TD link is like retrieving documents through a one-phrase query.



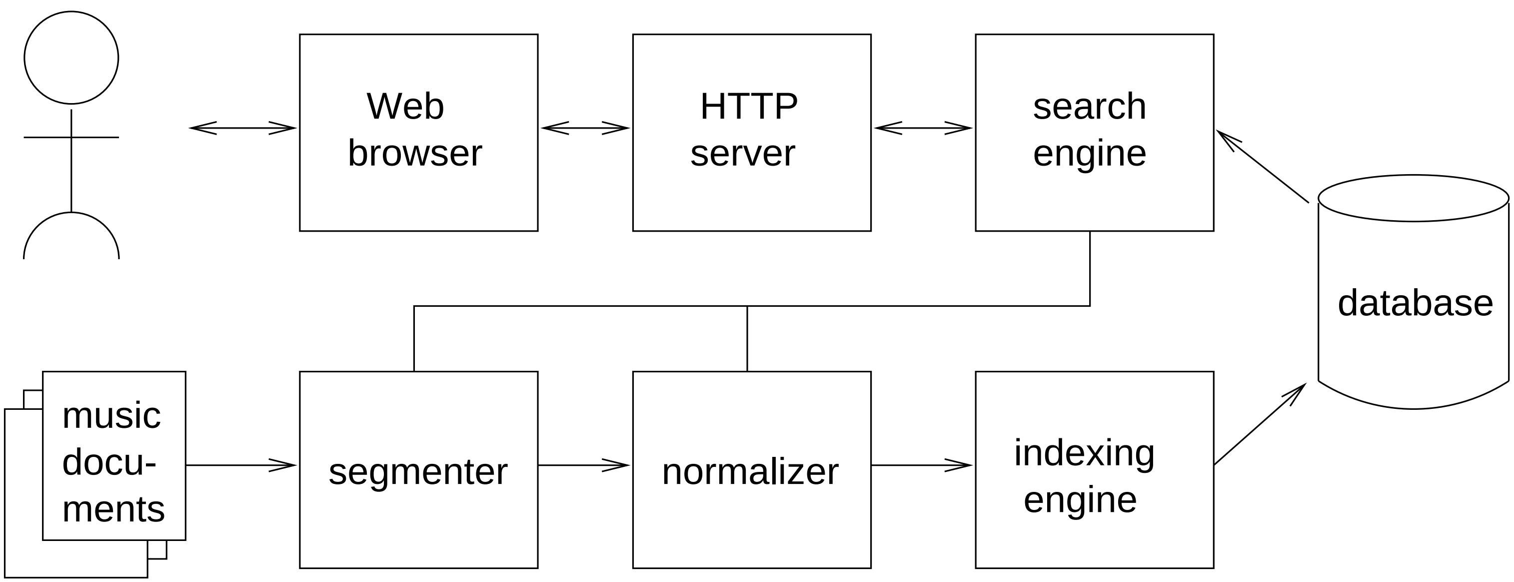
Hypermusic 5/6

The starting point (a) is a set of retrieved documents (black squares). The user follows a link from a retrieved documents to its similar ones (b). Then, she can follow links to phrases (c). From a phrase, she can follow links to similar phrases (d). Then, she follows links to new retrieved documents (e).



Hypermusic 6/6

The complete architecture may be depicted as



We developed a system for musical information retrieval called **SMILE.**

Smile 1/5

a **S**ystem for **M**usical **I**nformation retrieva**L** **E**nvironments

SMILE is available on the Internet at URL:

`http://livia.dei.unipd.it/smile)`

Smile 2/5

- SMILE implements the designed methodology to cope with content-based musical information retrieval.
- In SMILE content representation has been addressed both at *indexing* and *retrieval* stages.
- SMILE relies on fully automatic processes – segmentation, normalization, indexing, retrieval, and link construction.
- SMILE integrates *browsing* and *querying* music collections within a coherent methodological and implementational framework.

Smile 3/5

The SMILE prototype runs on a client-server architecture, where:

- the **client side** consists of a standard Java-enabled Web browser running a Java applet

- the **server side** consists of a musical indexing and search engines, the segmenter, the normalizer, other that the usual HTTP server

Smile 4/5: Client Side

The Java Applet

- the applet is used to create interactively the musical queries, and to tune the retrieval.

- a virtual keyboard can be played through the mouse, and queries can be created both recording a short music performance and inserting the notes one by one

- The applet converts the music score to a text notation that is used to invoke the server through a cgi-bin script. To hear the performance it is needed the JavaSound package that is available with Java version 2.1.3.

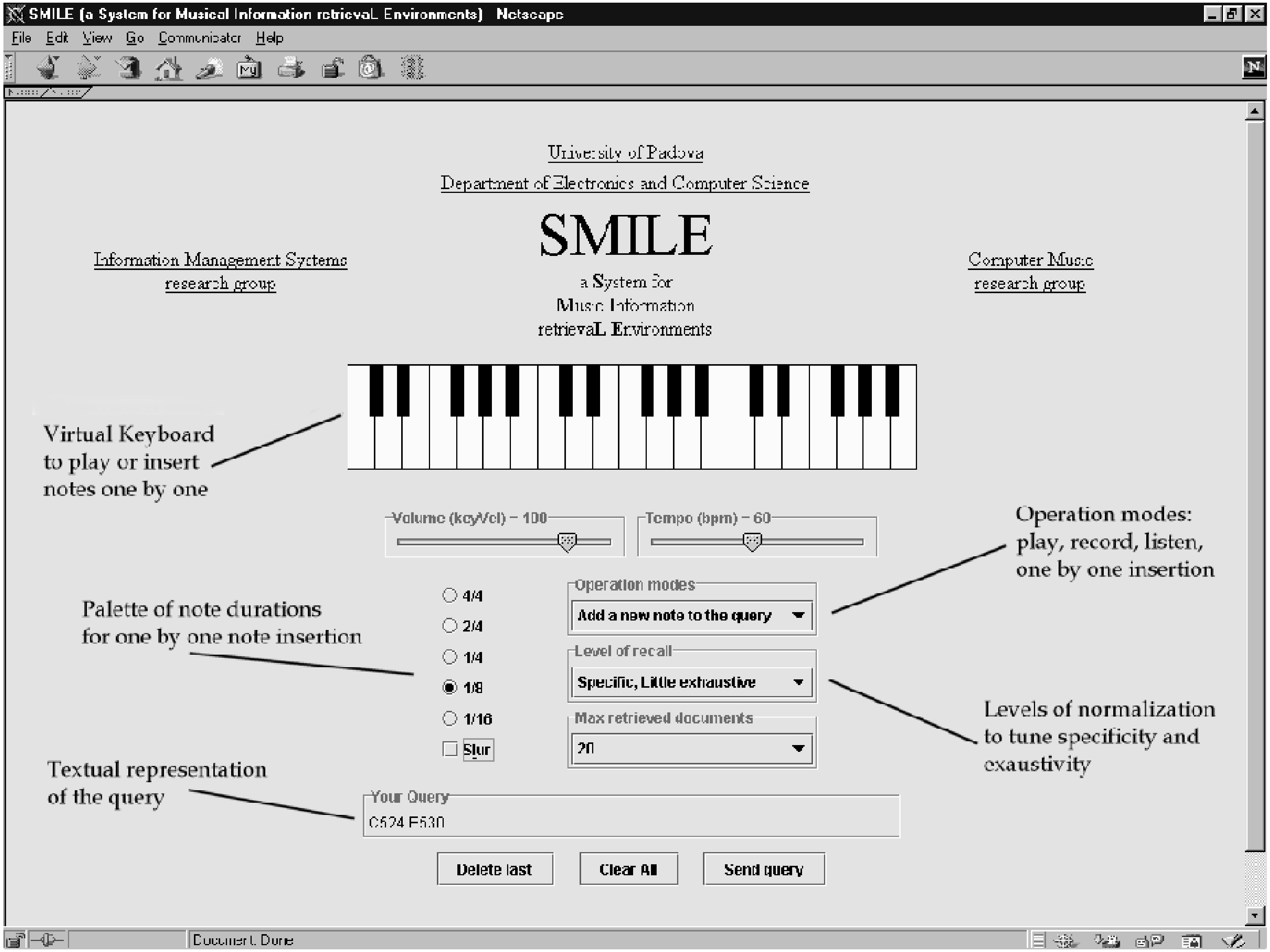
Smile 5/5: Server Side

- **The Segmenter**: takes an inner representation of a musical piece and creates a list of its musical phrases. The same routine is used to segment both the documents and the query.

- **The Normalizer**: creates four different representations, with an increasing level of generality, of the musical phrases produced by the segmenter.

- **The Indexing Engine**: takes the normalized textual musical scores as input and produces the indexes being used at search time.

- **The Search Engine**: takes the musical query being played using the Java applet, runs the segmenter and the normalizer taking the query as input, and accesses the indexes to retrieve the matching musical documents.


SMILE (a System for Musical Information retrievaL Environments) Netscape
File Edit View Go Communicator Help
University of Padova
Department of Electronics and Computer Science
SMILE
a System for
Music Information
retrievaL Environments
Information Management Systems
research group
Computer Music
research group
Virtual Keyboard
to play or insert
notes one by one
Volume (keyVel) – 100
Tempo (bpm) – 60
Operation modes:
play, record, listen,
one by one insertion
Palette of note durations
for one by one note insertion
4/4
2/4
1/4
1/8
1/16
Slur
Operation modes
Add a new note to the query
Level of recall
Specific, Little exhaustive
Max retrieved documents
20
Levels of normalization
to tune specificity and
exaustivity
Textual representation
of the query
Your Query
C524 F530
Delete last
Clear All
Send query
Document: Done

Query results - Netscape

File Edit View Go Communicator Help

Your original query was *64.48+65.48+67.48+71.48+72.48+74.48*

After segmentation and normalization the query is *O1P1P1 O1P1P1*

The system retrieved **10 Musical Documents**

1	0.458%	JSB0556.MID		Sound File	Related Musical Documents	Related Musical Phrases
2	0.4388%	WAM0331C.MID	Mozart - Piano Sonata #16 K331-c-RondoAllaTurca Midi by: B.Fisher	Sound File	Related Musical Documents	Related Musical Phrases
3	0.4165%	JSB1004D.MID	untitled Solo Violin ------------------------------------- Partita No. 2 in D minor for Solo Violin - BWV 1004 4th Movement: Gigue by Johann Sebastian Bach (1685-1750) Sequenced with Cakewalk Pro Audio by David J. Grossman (freely distributable) ------------------------------------- e-mail: JSBachBass@aol.com Dave's Home Page is at: http://pwp.starnetinc.com/daveg ------------------------------------- Original Filename: 4-vp2gig.mid Last Modified: 9-11-96	Sound File	Related Musical Documents	Related Musical Phrases
4	0.3903%	JSB0880B.MID	BWV 880/2	Sound File	Related Musical Documents	Related Musical Phrases
5	0.3829%	LVB033E.MID	Bagatelle (No.5 of "Seven Bagatelles", Opus 33) Allegro ma non troppo Ludwig van Beethoven Performed and sequenced by: Wesley Venable Instrument: Yamaha	Sound File	Related Musical Documents	Related Musical Phrases

Document: Done

Related Musical Documents - Netscape

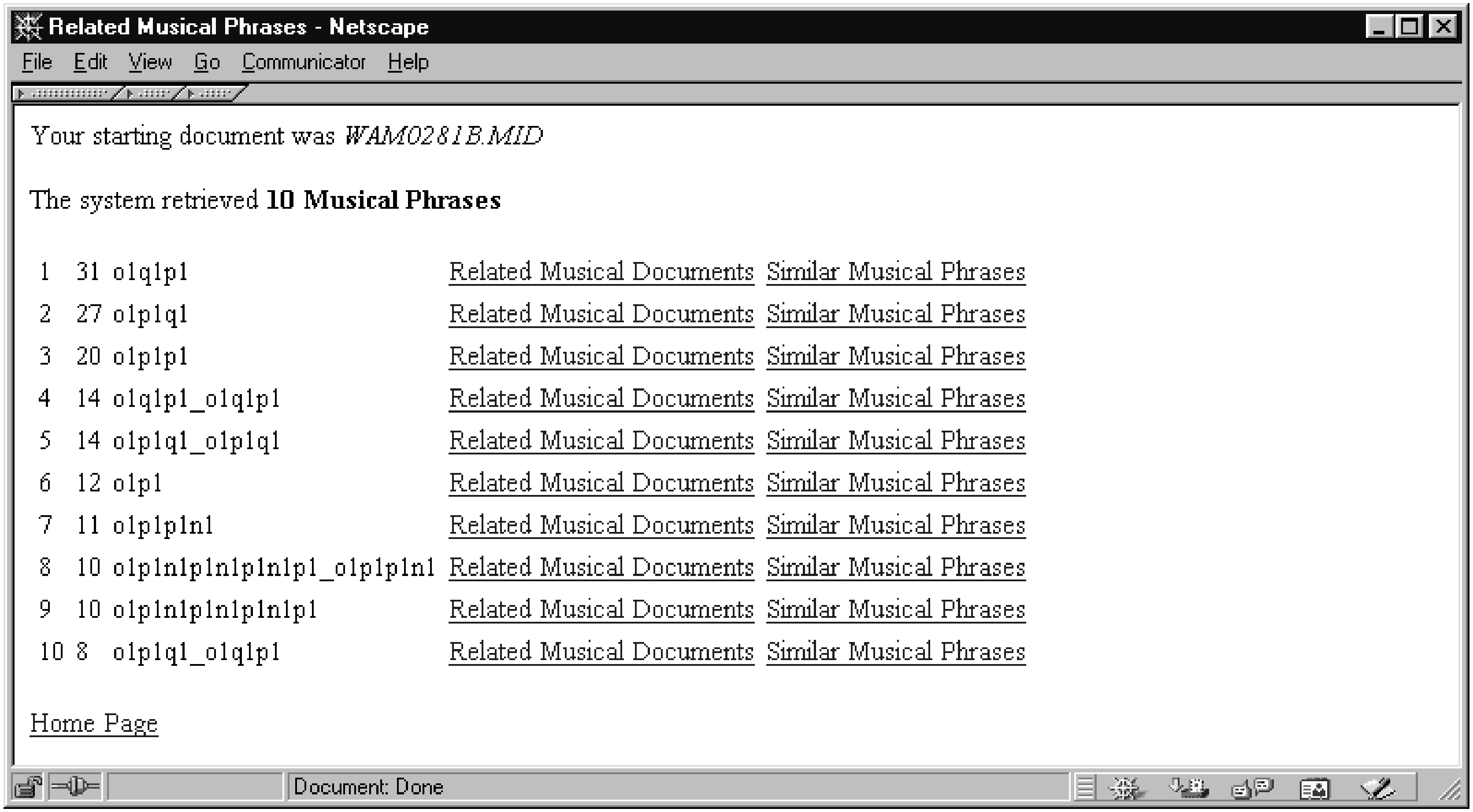
File Edit View Go Communicator Help

Your starting document was *JSB0556.MID*

The system retrieved **10 Musical Documents**

2	0.7214%	WAM0281B.MID	Mozart - Piano Sonata #6 K281 -b-Andante Midi by: B.Fisher	Sound File	Related Musical Documents	Related Musical Phrases
3	0.698%	WAM0357.MID	Primo Secondo Title: Sonate in Gmaj (K. 357, 1786) Composer: W. A. Mozart Sequence: Faren Raborn, Copyright 1996 Email: raborn@dtx.net	Sound File	Related Musical Documents	Related Musical Phrases
4	0.6889%	WAM0331A.MID	Mozart - Piano Sonata # 16 K331-a-AndanteGrazioso Midi by: B.Fisher	Sound File	Related Musical Documents	Related Musical Phrases
5	0.6371%	JSB0998A.MID	untitled David Grossman Solo Lute/Harpsichord Solo Lute/Harpsichord ------------------------------------ Prelude, Fugue, and Allegro in Eb major for Solo Lute - BWV 998 3rd Movement: Allegro by Johann Sebastian Bach (1685-1750) Sequenced with Cakewalk Pro Audio by David J. Grossman (freely distributable) ------------------------------------ e-mail: JSBachBass@aol.com Dave's Home Page is at: http://pwp.starnetinc.com/daveg ------------------------------------ Original Filename: pfa-3alg.mid Last Modified: 9-11-96	Sound File	Related Musical Documents	Related Musical Phrases
6	0.6327%	WAM0330.MID	Title: Sonata No. 10 in Cmaj (K. 330, 1778) Composer: W. A. Mozart Sequence: Faren Raborn, Copyright 1996 Email: raborn@dtx.net Site: www.dtx.net/~raborn	Sound File	Related Musical Documents	Related Musical Phrases

Document: Done


Related Musical Phrases - Netscape
File Edit View Go Communicator Help
Your starting document was *WAM0281B.MID*
The system retrieved **10 Musical Phrases**
1 31 o1q1p1 Related Musical Documents Similar Musical Phrases
2 27 o1p1q1 Related Musical Documents Similar Musical Phrases
3 20 o1p1p1 Related Musical Documents Similar Musical Phrases
4 14 o1q1p1_o1q1p1 Related Musical Documents Similar Musical Phrases
5 14 o1p1q1_o1p1q1 Related Musical Documents Similar Musical Phrases
6 12 o1p1 Related Musical Documents Similar Musical Phrases
7 11 o1p1p1n1 Related Musical Documents Similar Musical Phrases
8 10 o1p1n1p1n1p1n1p1_o1p1p1n1 Related Musical Documents Similar Musical Phrases
9 10 o1p1n1p1n1p1n1p1 Related Musical Documents Similar Musical Phrases
10 8 o1p1q1_o1q1p1 Related Musical Documents Similar Musical Phrases
Home Page
Document: Done

Related Musical Phrases - Netscape

File Edit View Go Communicator Help

Your starting term was *o1p1p1n1*

The system retrieved **10 Musical Phrases**

1	20.4%	o1p1q1n1	Related Musical Documents	Related Musical Phrases
2	18.66%	o1o1o1o1p1	Related Musical Documents	Related Musical Phrases
3	18.64%	o1p1p1_o1p1n1	Related Musical Documents	Related Musical Phrases
4	18.62%	o1n1n1_o2n1	Related Musical Documents	Related Musical Phrases
5	18.62%	o1p1p1n1_o1p1p1p1	Related Musical Documents	Related Musical Phrases
6	18.62%	o1p1p1p1m2	Related Musical Documents	Related Musical Phrases
7	18.62%	o1q1_o1n1p1p1	Related Musical Documents	Related Musical Phrases
8	18.26%	o1p1p1_o1n1p1p1	Related Musical Documents	Related Musical Phrases
9	17.72%	o1n2_o1n1	Related Musical Documents	Related Musical Phrases
10	17.67%	o1o1_o1p1p1n1	Related Musical Documents	Related Musical Phrases

Document: Done

Related Musical Documents - Netscape

File Edit View Go Communicator Help

Your starting term was *olplqlnl*

The system retrieved **10 Musical Documents**

1	2	untitled J. S. Bach Sonata for Flute and Harpsichord in B minor BWV 1030 Flute Harpsichord RH Harpsichord LH	Sound File	Related Musical Documents	Related Musical Phrases
2	2	OBOE HARPSICHORD HARPSICHORD Bach: Trio Sonata #1 Mvt. 1 Seq. by:	Sound File	Related Musical Documents	Related Musical Phrases
3	1		Sound File	Related Musical Documents	Related Musical Phrases
4	1	Piano dx Pianp sin Violin I Violin 2 Stråkbas Info J.S Bach BWV 1055, 1st movement, sequenced 1997 by Per Gustavsson J.S Bach: BWV 1055, 1st movement, sequenced 1997 by Per Gustavsson solo viol violin 1 violin 2 viola cello basso con Bach's violin works(all 5concerto 6 v-p son 6 sonatas partitas for solo violin) available as M-file Dietrich Gewissler 74030 .2433@ compu serve.com\	Sound File	Related Musical Documents	Related Musical Phrases
5	1	soprano alto tenor bass accomp treble accomp bass 12 - Chorus Messiah by Handel sequence by Michael Petri	Sound File	Related Musical Documents	Related Musical Phrases
6	1	Mozart - Piano Concerto #15 K450 -b -Andante Midi by: B.Fisher	Sound File	Related Musical Documents	Related Musical Phrases
7	1	Primo Secondo Title: Sonate in Fmaj (K. 497, 1786) Composer: W. A. Mozart Sequence: Faren Raborn, Copyright 1996 Email: raborn@dtx.net	Sound File	Related Musical Documents	Related Musical Phrases
8	1	1812 Overture Flutes Oboes Clarinets English Horn Bassoons French Horns Cornets Trumpets Trombones & Tuba Tympani Snare Drum Bass Drum Cannon Triangle & Bells Cymbals Guns Violins Violas Cellos Contrabass 1812 OVERTURE by P. I. Tschaikovsky Adapted and Sequenced by Robert C. Goodyear	Sound File	Related Musical Documents	Related Musical Phrases

Document: Done

Conclusions 1/3

- A new methodology for musical information retrieval has been developed.

- SMILE, a prototype system based on this methodology, is available on the Web.

- The approach can be integrated with fixed field-based search tools.

Conclusions 2/3: Pros

- Automatic segmentation can help decide how documents can be segmented and indexed.
- Normalization allows to tune specificity and exhaustivity of retrieval.
- Results can be ranked by a Retrieval Status Value and thresholds can be applied.

Conclusions 3/3: Cons

- The algorithms have been tested only for Western music (Baroque and Romantic periods).

- It is likely that inter-manual segmentation is inconsistent, then automatic segmentation must be evaluated.

- No information about harmony is employed.

- It is *unclear whether musical phrases has meaning* as textual words do.

Future Work

- An evaluation of the segmentation algorithm is in progress. Melomane, musicians, scholars, and students are involved.

- Integration of the system inside the Digital Library for the Venetian Music.

- A prototype system storing musical documents with different formats (e.g. audio files, digitized music scores, MP3, MIDI) will be developed.