Automatic Data Understanding

the Tool for Intelligent Man-Machine Communication

Władysław Homenda

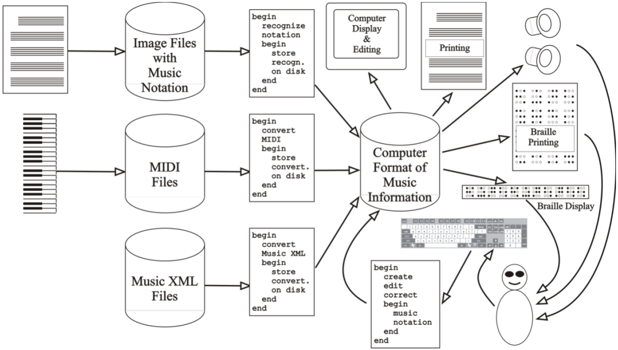
Faculty of Mathematics and Information Science, Warsaw University of Technology
Faculty of Economics and Informatics in Vilnius, University of Bialystok

SEMAPRO, Barcelona, November 12th, 2017

Outline

Motivation


BRAILLE
Main page
Potential
Results
Library
Apropos
Login
User
Password
Login
Register
Forgotten your password?
Homepage
This page is dedicated to project Removing barriers to functioning of people with disabilities in information society - design and implementation of computer system aimed at processing music information for blind people. carried out by Systems Research Institute of Polish Academy of Sciences between 2009-2012. The project was funded by The National Centre For Research and Development. Scientific publications, conference presentations and prototype music information processing program for blind people were published during project implementation.
Usuwanie barier funkcjonowania osób niepełnosprawnych w społeczeństwie informacyjnym.
Braille Score - projekt i wykonanie komputerowego systemu przetwarzania informacji muzycznej dla osób niewidomych.
Projekt rozwojowy NCBiR nr N R02 0019 06/2009
realizowany w Instytucie Badań Systemowych
Polskiej Akademii Nauk
w Warszawie,
2009 – 2012



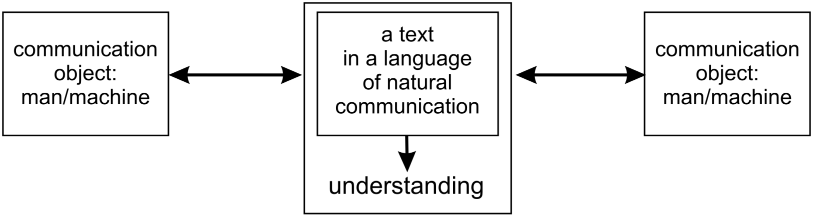
Braille Score is a prototype computer system to process music information dedicated to both sighted and blind people, cf. **http://braillescore.ibspan.waw.pl/**:

- create new information from scratch (e.g. own music record)
- edit existing music information
- multiway conversion between following formats
  - printed paginated music notation
  - Music XML
  - Braille Music
  - MIDI

Communication

- Communication is a process of information exchange
- Communication is accomplished in some language
  - composer communicates musicians how to perform in a given key, he uses a key signature, the signature is put in respective place of music notation
  - teacher instructs students how to solve a problem, he instructs in a natural language
- Information being exchanged is expressed and/or described as a text in some language
- Intelligent communication is just exchanging texts, which express information and/or describe data
- Texts being exchanges should be understood in communication process

Communication



An overall diagram of communication.
Either man or machine is considered as object of communication.

Languages - tools of communication

Communication between human beings is carried out in languages

- natural languages, like English, Spanish, Russian
- music notation, gesture languages
- C/C++/C# programming languages
- we can also consider painting, sculpture, music etc. as tools of communication

Ad hoc taxonomy

- languages of natural communication
- languages of formal communication
- languages of arts

Languages of natural communication

The common features of ***languages of natural communication***:

- created, developed and used prior to their formal codification and
- no full and strict formal definition

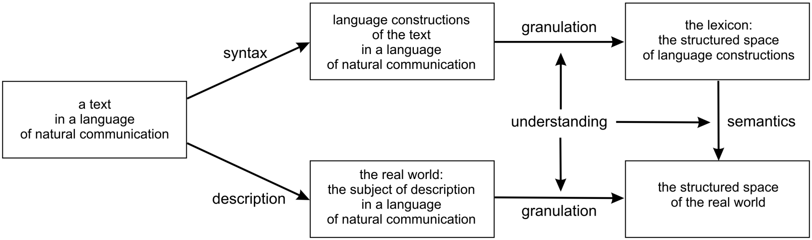
Moreover:

- they are used intuitively with high-level of flexibility
- some kind of anarchy, mistakes and errors in usage of languages of natural communication does not break communication

Examples of languages of natural communication:

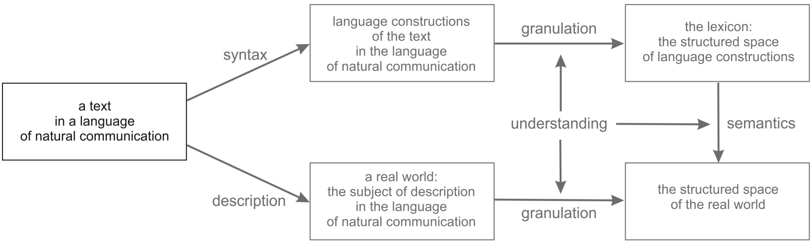
- natural languages
- music notations
- gesture languages

Information understanding



An overall diagram of the paradigm of information understanding in context of communication

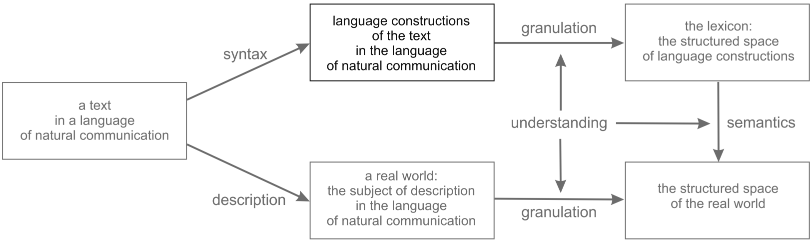
Text in a language



Examples of texts in languages of natural communication:

- texts in a natural languages
- scores of paginated music notation
- scores of Braille music notation

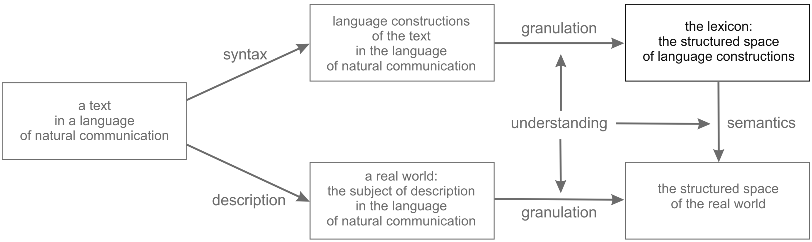
Language constructions



Language constructions:

- texts, sentences and phrases in a natural language
- scores of paginated music notation, parts and phrases of scores

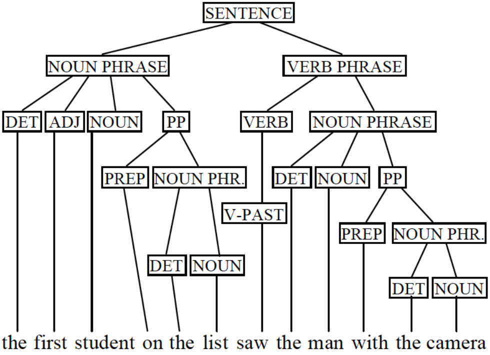
Lexicon



Lexicon:

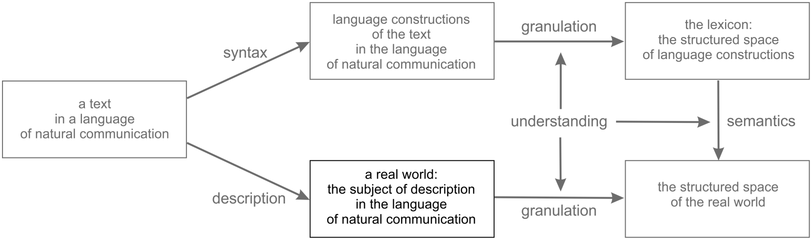
- a structured space of language constructions
- each construction supplemented with a derivation (parsing) subtrees
- the lexicon includes relations between its items

Natural language constructions - derivation (parsing) trees



Bargiela A., Homenda W.: Information structuring in natural language communication: Syntactical approach, Journal of Intelligent and Fuzzy Systems, 17(6), 575-581, 2006

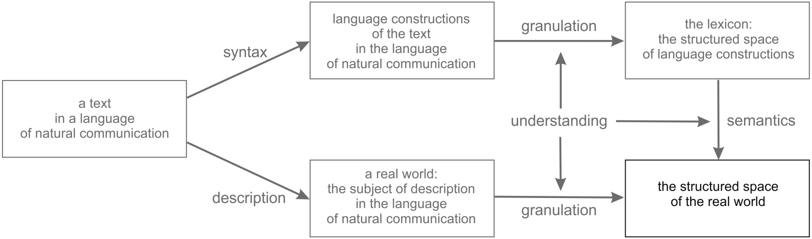
Real world



Real world:

- the subject of descriptions in a language of natural communication
- a world of things, sensations, thoughts, ideas etc.
- a part of this world is described by a text in a language of natural communication

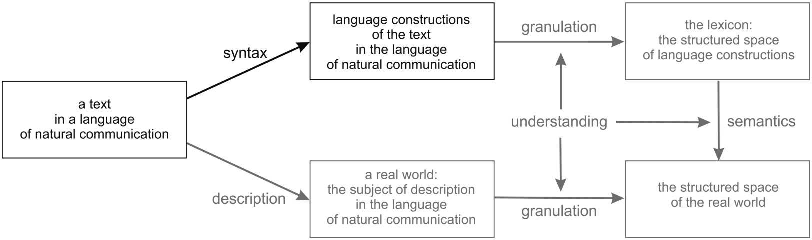
Structured space of a real world



Structured space of a real world:

- the space of things sensations, thoughts, ideas etc.
- provided with structures based on relations between items and sets of items

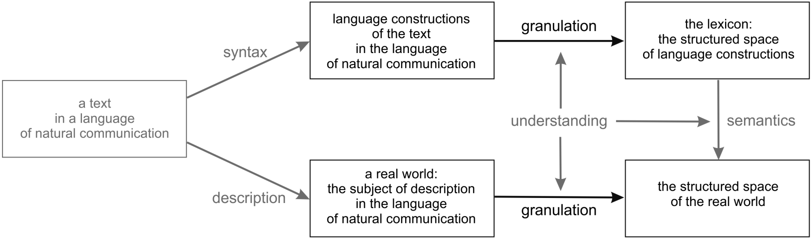
Syntax



Syntax, syntactic structuring:

- concerns texts of languages of natural communication
- the study on how words fit together structures describing real worlds
- usually described by a kind of a grammar
- our study is based on context-free grammars

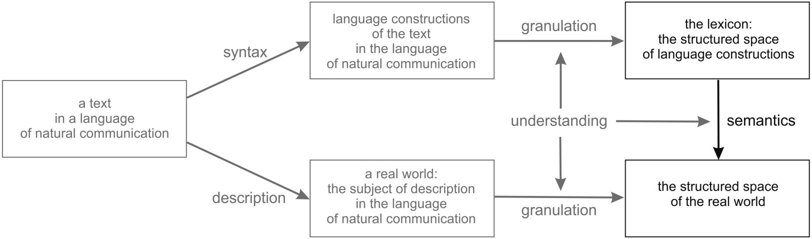
Granulation



Granulation

- a concept of structuring
  - the space of language constructions
  - the real world
- creation of the lexicon on the basis of language constructions is an apparent application of the paradigm of granularity

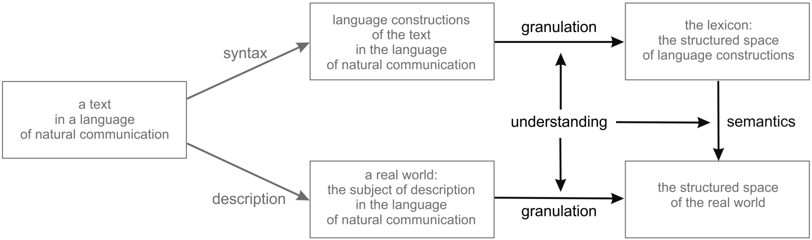
Semantics



Semantics:

- captures a relation between items of lexicon and items of the structured space of real world
- in other words, semantics casts the lexicon onto the structured space of the real world
- in general, such casts/relations are many-to-many dependencies
- it may be one-to-many, many-to-one or, rarely, one-to-one relation

Understanding



Understanding is an ability:

- to construct granular structures in both spaces
  - lexicon
  - real world
- and to construct semantics

Generalities

- understanding of data is a fundamental feature of communication
- automatic understanding of data is a fundamental feature of intelligent human-machine communication
- the paradigm of automatic data understanding is the universal approach to intelligent human-machine communication
- the paradigm of automatic data understanding involves syntactic description, semantic analysis and granular structuring as general methods

Case studies

- developing detailed methods of automatic data understanding, which would be domain independent, is not possible at the current level of research and technology
  - we do not attempt to generalize these methods in a way that they would be directly applicable in any subject of communication
  - instead, we provide a case study in chosen domains
  - details discussed in this presentation may not be universally applicable in any domain
- we analyze structured spaces of music data oriented to paginated music notation.
- te discussion is focused on operations performed on spaces of music information.

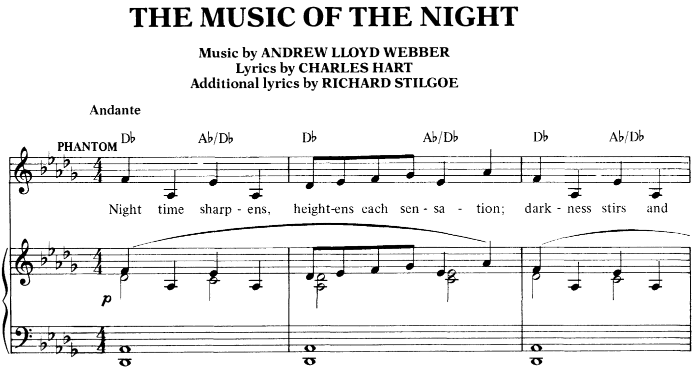
Example 1


SUITE NO. 3 IN D
J. S. BACH
arranged by Thomas Arnold Johnson
Adagio
pp
p
mf

Example 2


10. Polonaise
(from Sonata for Flute and Piano)
LUDWIG VAN BEETHOVEN (1770-1827)
Moderato
mf
f
mf
f
p
p

Example 3


THE MUSIC OF THE NIGHT
Music by ANDREW LLOYD WEBBER
Lyrics by CHARLES HART
Additional lyrics by RICHARD STILGOE
Andante
PHANTOM
D♭ A♭/D♭ D♭ A♭/D♭ D♭ A♭/D♭
Night time sharp - ens, height-ens each sen - sa - tion; dark - ness stirs and
p

Grammars

$$G=(V,T,P,S)$$

- V is a finite set of *variables* (called also *nonterminals*)
- T is a finite set of terminal symbols (simply called *terminals*)
- a nonterminal S is the initial symbol of the grammar and
- P is a finite set of productions, $\alpha \to \beta$

A derivation in a grammar is a finite sequence of strings of nonterminals and terminals such that:

- the first string in this sequence is just the initial symbol of the grammar and
- for any two consecutive strings in the sequence, the later one is obtained from the former one applying a production, i.e. by replacing a substring of the former string equal to the left hand side of the production with the right hand side of the production

Chomsky hierarchy

$$\text{RgL} \subsetneq \text{CFL} \subsetneq \text{CSL} \subsetneq \text{REL} \subsetneq \text{ALL}$$

- Chomsky's taxonomy of languages of natural communication is not known
- each language of natural communication should be individually classified
- it is even not known, if natural languages are context-free or context-sensitive
- many attempts to demonstrate that English is a context-sensitive language and not context-free one have occurred either to be unsuccessful or incorrect, cf. Jurafsky D., Martin J. H., Speech and Language Processing, Prentice Hall Inc., 2000.

Chomsky hierarchy

$$\text{RgL} \subsetneq \text{CFL} \subsetneq \text{CSL} \subsetneq \text{REL} \subsetneq \text{ALL}$$

With regard to paginated music notation

- it seems to be context-sensitive language
- it contains repetitions (i.e. constructions of the forms $ww, www, wwww, \ldots$ such that w is a part of paginated music notation), which are typical context sensitive structures
- in practice, length/cardinality of languages of natural communication would be limited by a constant value, therefore they could be considered as context-free (even regular)

Context-free grammars

$$G=(V,T,P,S)$$

Context-free grammars:

- productions are of the form $A \rightarrow \beta$
- well elaborated in theory and practice
- easy applicable

For a context-free grammar a derivation can be outlined in a form of derivation (parsing) tree, i.e.

- the root of the tree is labelled with the initial symbol of the grammar and
- for any internal vertex labelled by the left side of a production, its children are labelled by symbols of the right side of the production

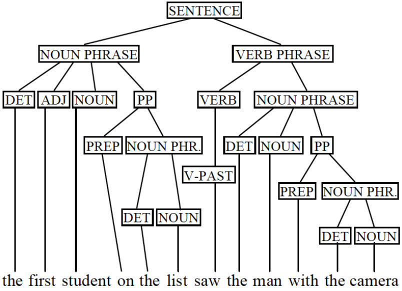
Natural language constructions - derivation (parsing) trees

<SENTENCE>	→	<NOUN PHRASE> <VERB PHRASE>
<NOUN PHRASE>	→	<DET> <ADJ> <NOUN> <PP>
	→	<DET> <NOUN>
<VERB PHRASE>	→	<VERB> <NOUN PHRASE>
	→	<VERB>
<PP>	→	<PREP> <NOUN PHRASE>
	→	<VERB>

. . .

<DET>	→	the
	→	a
	→	ε
<NOUN>	→	student
	→	list
	→	man
	→	camera

. . .



Syntactic approach

Crucial stage in the wide spectrum of tasks, for instance:

- translating programming languages
- processing natural languages
- processing paginated music notation

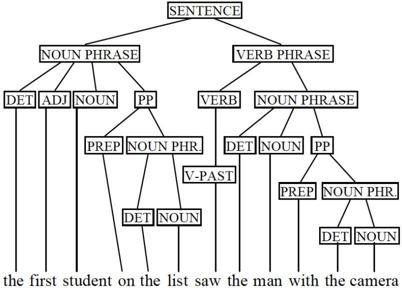
Syntactic approach - the following tools used in processing languages

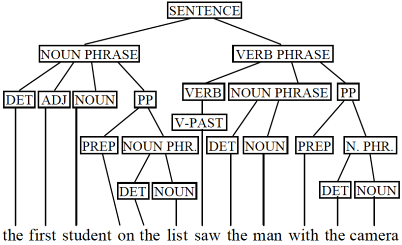
- grammars
- automata

Syntactic approach

- effective
- not precise enough

Natural language constructions - derivation (parsing) trees
Ambiguities of syntactic structuring





Bargiela A., Homenda W.: Information structuring in natural language communication: Syntactical approach, Journal of Intelligent and Fuzzy Systems, 17(6), 575-581, 2006

Context-free syntactic descriptions

- the Chomsky's taxonomy of languages of natural communication is of less importance for practical applications
- even if context-free, they are too complex to be fully formalized,
- despite a definite set of rules defining languages of natural communication
  - the rules are highly complicated
  - the number of rules is huge
- such rules can be broken with little impact on communication
- thus, a description of a language of natural communication must definitely be highly flexible and deeply tolerant to natural anarchy of its subjects.
- pure context-free tools are inadequate for successful processing

Context-free syntactic descriptions

- the proposed approach to describing languages of natural communication will rely on sensible application of context-free methods:
  - they are applied locally in order to cover a usually very limited subset of language constructions
  - they cover whole languages, i.e. they generate all constructions of a language, but also constructions not belonging to it
- we assume that the context-free methods will not be applied unfairly to generate incorrect constructions or constructions not belonging to languages being described.
- these assumptions allow only for a raw approximation of languages of natural communication.

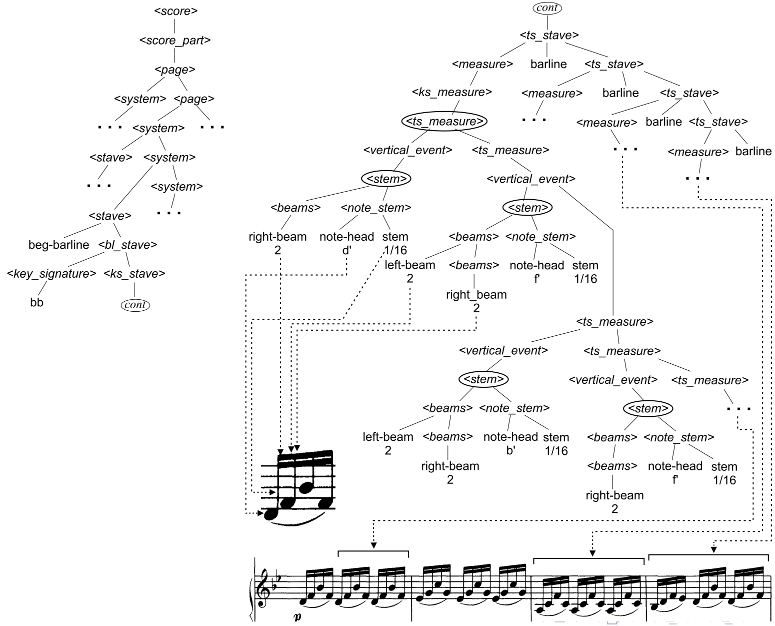
Context-free syntactic descriptions

- a simplified context-free grammars are used for the purpose of syntactical structuring of paginated music notation
- the grammars will be applied to analyze constructions, which are assumed to be well grounded pieces of paginated music notation
- such grammars can neither be applied for checking correctness of constructions of paginated music notation, nor to generate such constructions.

Context-free descriptions

<score> → <score_part> <score> | <score_part>
<score_part> → <page> <score_part> | <page>
<page> → <system> <page> | <system>
<system> → <stave> <system> | <stave>
<system> → <part_name> <stave> <system> | <part_name> <stave>
<part_name> → Flute | Piano | *etc.*
<stave> → beginning-barline <bl_stave> | <clef> <bl_stave>
<stave> → beginning-barline <clef> <bl_stave> | <bl_stave>
<bl_stave> → <key_signature> <ks_stave> | <ks_stave>
<ks_stave> → <time_signature> <ts_stave> | <ts_stave>
<ts_stave> → <measure> <barline> <ts_stave> | <measure> <barline>
<measure> → <change_of_k_sign.> <ks_measure> | <ks_measure>
<ks_measure> → <change_of_t_sign.> <ts_measure> | <ts_measure>
<ts_measure> → <vertical event> <ts_measure> | <vertical event>
<vertical event> → <stem> <vertical event> | <stem>
<stem> → <beams> <note_stem> | <flags> <note_stem> | <note_stem>
<beams> → left-beam | right-beam | left-beam right-beam
<clef> → treble_clef | bass_clef
<flags> → flag <flags> | flag
<note_stem> → note-head <note_stem> | note-head

Context-free descriptions



<score> → <score_part> <score> | <score_part>
<score_part> → <page> <score_part> | <page>
<page> → <system> <page> | <system>
<system> → <stave> <system> | <stave>
<system> → <part_name> <stave> <system> | <part_name> <stave>

<system> → *<stave> <system>* | *<stave>*

<system> → *<part_name> <stave> <system>* | *<part_name> <stave>*

<part_name> → Flute | Piano | *etc.*

<stave> → beginning-barline *<bl_stave>* | *<clef> <bl_stave>*

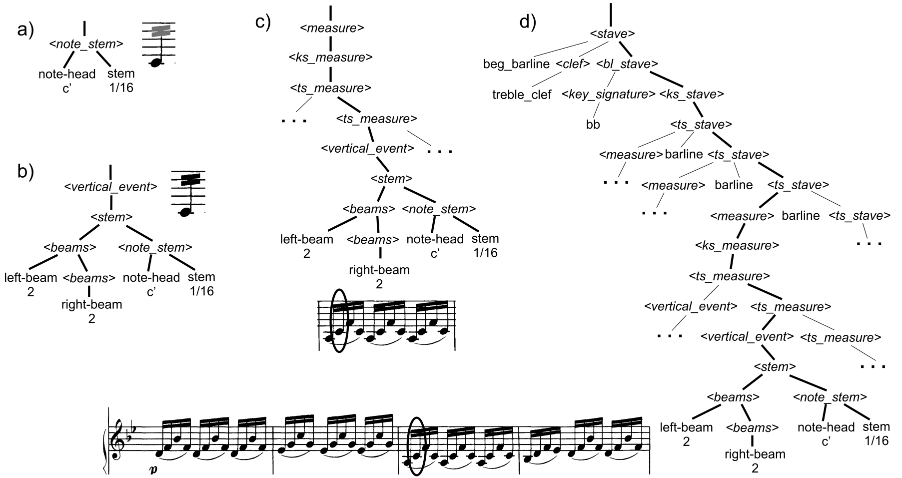
<stave> → beginning-barline *<clef> <bl_stave>* | *<bl_stave>*

<bl_stave> → *<key_signature> <ks_stave>* | *<ks_stave>*

<ks_stave> → *<time_signature> <ts_stave>* | *<ts_stave>*

<ts_stave> → *<measure> <barline> <ts_stave>* | *<measure> <barline>*

The Lexicon Polonaise from Sonata for Flute and Piano by Beethoven



Valuation relation

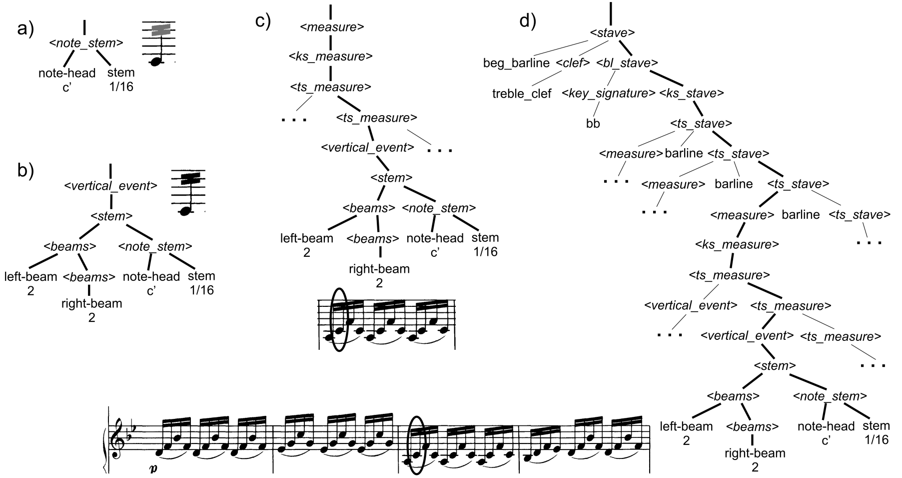
- lexicon elements describe objects in a real world
- in our case the real world includes sounds produced when a piece of music is being performed
- for the sake of simplicity we assume that sounds are of the form of notes described by three parameters: pitch, beginning time and duration of a note
- a triple roughly description of a sound/note can be supplemented by other features, e.g. volume (loudness), articulation (legato, staccato, portato) etc.

Semantics is a relation V in the lexicon L and the world W, called *valuation relation*:

$$V \subset L \times W$$

Valuation relation J. S. Bach, Suite (Ouverture) No. 3 in D major, BWV 1068

Valuation relation Polonaise from Sonata for Flute and Piano by Beethoven



Granular computing paradigm

- raises from a need to comprehend the problem and provide a better insight into its essence rather than get buried in all unnecessary details
- in this sense, granulation serves as an abstraction mechanism that reduces an entire conceptual burden
- as a matter of fact, by changing the size of the information granules, we can hide or reveal a certain amount of details to deal with during a certain design phase
- it is worth stressing that information granules
  - not only support conversion of clouds of detailed data into more tangible information granules
  - but, very importantly, afford a vehicle of abstraction that allows to think about granules as different conceptual entities.

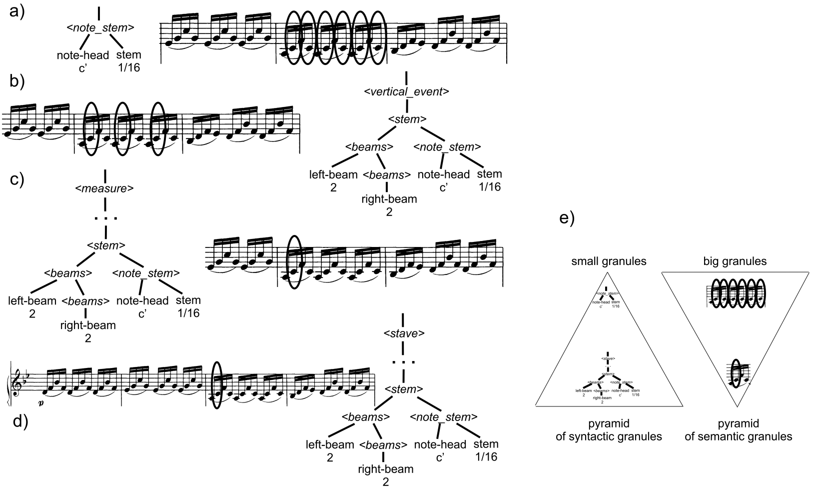
Granular computing paradigm

- utilizing granular computing to data processing
  - supports conversion of clouds of detailed data into more tangible information granules and
  - provides a vehicle of abstraction that allows to think of granules as different conceptual entities.
- summarizing
  - a collection of data items is subjected to a process of mining conceptual entities of information,
  - this process finally leads to an overall understanding of the processed subject
  - the process of structuring the space of information fits perfectly the paradigms of granular computing and information granulation

Granular structuring of music information

- the description of music notation as well as music notation itself could be innately subjected to the paradigm of granular computing elucidation
- granular computing as opposed to numeric computing is knowledge-oriented; information granules exhibit different levels of knowledge abstraction, what strictly corresponds to different levels of granularity
- depending upon the problem at hand, we usually group granules of similar size (i.e. similar granularity) together into a single layer
- in total, the arrangement of this nature gives rise to the information pyramid
- information granularity implies the usage of various techniques that are relevant for the specific level of granularity

Forming syntactic and semantic granules



Understanding

- data understanding is an ability to recognize semantics and granulation of information
- we do not attempt to develop a theoretical framework for automatic data understanding
- we believe that case studies will be more informative in context of this presentation
- we discuss examples of structural operations in general and specifically in the space of music information
- such operations are the best manifestation of what we mean as data understanding

Selecting

- among structural operations selection is the very fundamental one
- selection is the basis for other common structural operations like copying, finding, replacing etc.
- technically, drawing a rectangle with a mouse is the most popular methods of selection on a screen
- possibly a multi selection can be done, usually with mouse moves or clicks while the key Ctrl is pressed
- more sophisticated selection targets are defined with UI tools: like menus, dialog boxes etc.
  - menus
  - dialog boxes
  - regular expressions

Selecting, copying and pasting in PSP - no understanding

Kindly assure that the Contact Volume Editor is given the name and email address of the contact author for your paper. The Contact Volume Editor uses these details to compile a list for our production department at SPS in India. Once the files have been worked upon, SPS sends a copy of the final pdf of each paper to its contact author. The contact author is asked to check through the final pdf to make sure that no errors have crept in during the transfer or preparation of the files. This should not be seen as an opportunity to update or copyedit the papers, which is not possible due to time constraints. Only errors introduced during the preparation of the files will be corrected.

Kindly assure that the Contact Volume Editor is given the name and email address of the contact author for your paper. The Contact Volume Editor uses these details to compile a list for our production department at SPS in India. Once the files have been worked upon, SPS sends a copy of the final pdf of each paper to its contact author. The contact author is asked to check through the final pdf to make sure that no errors
Kindly assure that the Contact r or preparation of then the name and email address of
the contact author for your pappyedit the papers, wlume Editor uses these details to
compile a list for our productioduced during the pin India. Once the files have been
corrected. each paper to its contact author.

Selecting, copyingand pasting in Word - no understanding

This round of checking takes place about two weeks after the files have been sent to the Editorial by the Contact Volume Editor, i.e., roughly seven weeks before the start of the conference for conference proceedings, or seven weeks before the volume leaves the printer's, for post-proceedings. If SPS does not receive a reply from a particular contact author, within the timeframe given, then it is presumed that the author has found no errors in the paper.

This round of checking takes place about two weeks after the files have been sent to the Editorial by the Contact Volume Editor, i.e., roughly seven weeks before the start of the conference for conference proceedings, or seven weeks before the volume leaves the printer's, for post-proceedings. This round

takes

about two weeks

for post-proceedings.

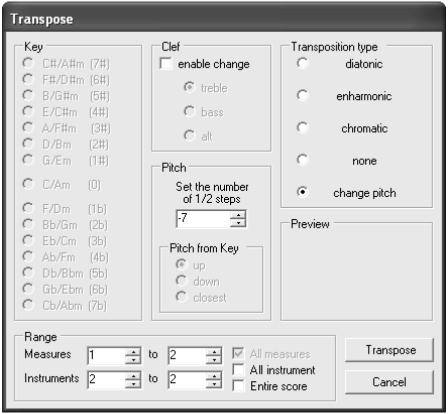
If SPS does not receive a reply from a particular contact author, within the timeframe given, then it is presumed that the author has found no errors in the paper.

Selecting data, J. S. Bach, Suite (Ouverture) No. 3 in D major, BWV 1068

An example of a raster image selection: a rectangular region of the screen is selected and highlighted. This selection is interpreted as a part of a raster image and not the second and third measures of the bottom stave of the paginated music notation. In this interpretation content of the image is of minor importance: it may be a part of printed music notation as well as a part of city map or a part of any other raster image.

Selecting

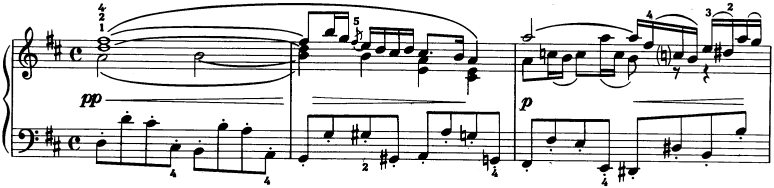
- selection is also the basis for domain specific operations
- such domain specific operations depend on selection:
  - voice line identification
  - transposition
  - conversion
  - harmonization
- domain dependent selections require advanced methods and tools to defined their targets



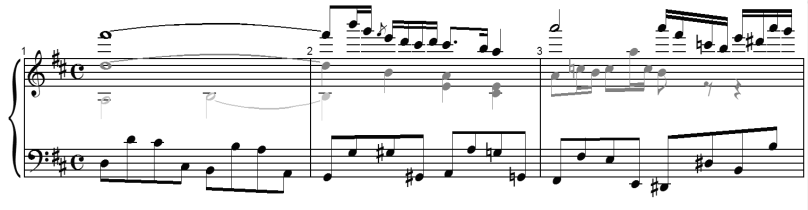
Searching Polonaise from Sonata for Flute and Piano by Beethoven

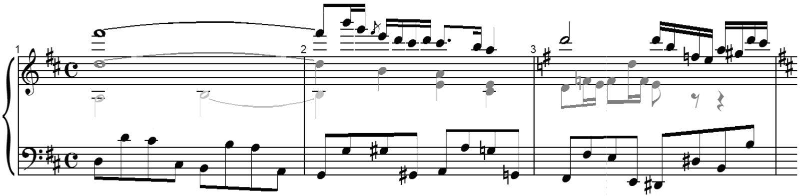
Illustration of searching. The pattern is in dotted ellipse, instances are dashed around. The upper part illustrates searching for exact matches (the same pitches, durations and time intervals between notes) without specific placement in the measure. The bottom part shows instances preserving pitch intervals, time intervals, durations and placement in the measure.

Identifying voice lines J. S. Bach, Suite (Ouverture) No. 3 in D major, BWV 1068

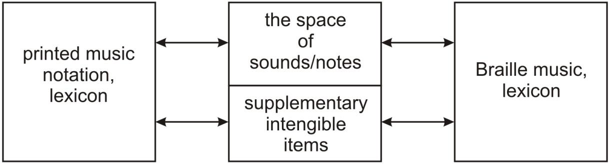


Operations





Converting to Braille



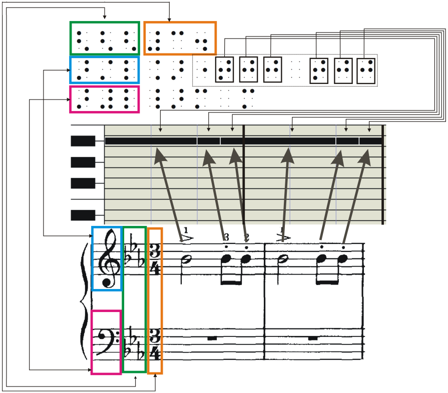
Conversion from printed music to a piece of Braille music is a mapping C:

$$C : L \to B$$

where: L and B are lexicons. However, automatic construction of such a mapping is extremely hard. The conversion was done as mappings

$$L \xrightarrow{V} S \xrightarrow{W^{-1}} B \qquad B \xrightarrow{W} S \xrightarrow{V^{-1}} L$$

Converting to Braille

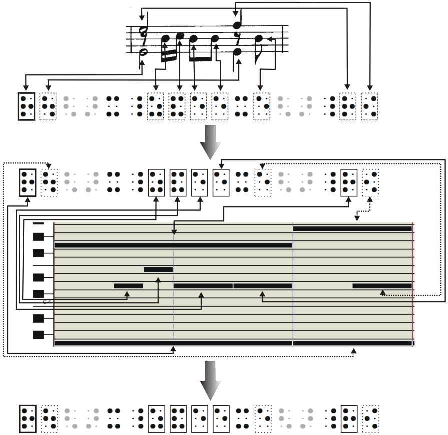


Valuation function: mappings from printed music notation and Braille music into the space of sounds. Derivation trees are not shown. The played notes are displayed as strips on the scale of piano keyboard.

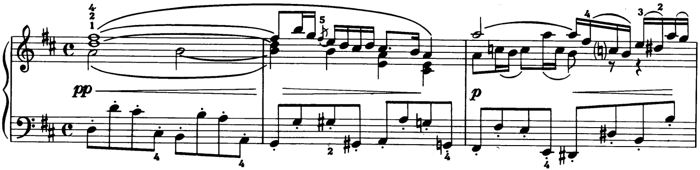
Processing Braille, hard operations

- assigning key (when missing depends only on notes pitches), needs calculation of pitches for all notes,
- searching for note or set of notes occurrence, needs calculation of duration for all notes and checking their relative location,
- searching for the first occurrence of a note with the specified pitch,
- searching for voices or voices' parts with the same rhythmic,
- searching for all notes that are played within specified note,
- raising/lowering specified note by specified number of halftones, needs calculation of pitch for the note and decision if an octave sign is required,
- score's part transposition or whole score transposition by any interval, needs calculation of exact pitch for all notes and decision if an octave sign is required

Converting to Braille



Dimensionality J. S. Bach, Suite (Ouverture) No. 3 in D major, BWV 1068



Time prior to pitch

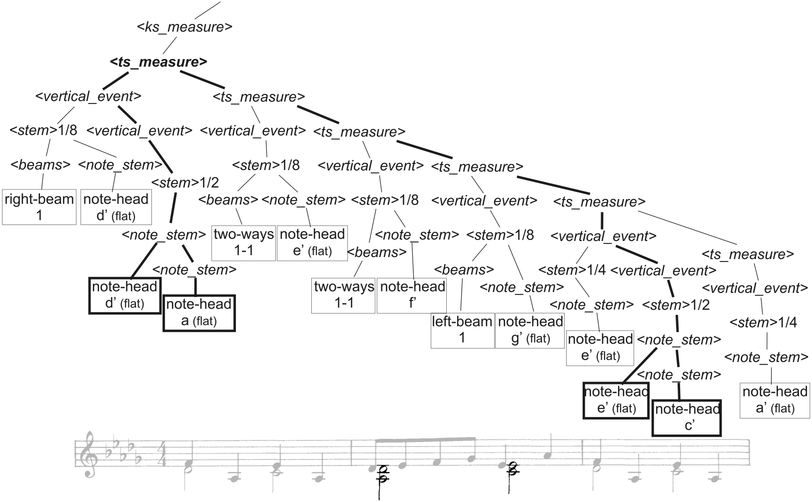
<stave> → beg-barline *<bl_stave>* | *<bl_stave>*
<bl_stave> → *<clef>* *<cl_stave>* | *<cl_stave>*
<cl_stave> → *<key_signature>* *<ks_stave>* | *<ks_stave>*
<ks_stave> → *<time_signature>* *<ts_stave>* | *<ts_stave>*
<ts_stave> → *<measure>* barline *<ts_stave>* | *<measure>* barline | *<measure>*
<measure> → *<change_of_k_sign.>* *<ks_measure>* | *<ks_measure>*
<ks_measure> → *<change_of_t_sign.>* *<ts_measure>* | *<ts_measure>*
<ts_measure> → *<vertical event>* *<ts_measure>* | *<vertical event>*

<vertical event> → *<stem>* *<vertical event>*
→ *<stem>*

<stem> → *<beams>* *<note_stem>*
→ *<flags>* *<note_stem>*
→ *<note_stem>*

<beams> → left-beam *<beams>* | right-beam *<beams>*
<beams> → left-beam | right-beam | two-ways
<flags> → flag *<flags>* | flag
<note_stem> → note-head *<note_stem>* | note-head

Time prior to pitch



Pitch prior to time

Time prior to pitch orientation forced by the following mechanism:

$$
\begin{aligned}
<ts\_measure> &\rightarrow <vertical\ event>\ <ts\_measure> \\
&\rightarrow <vertical\ event> \\
<vertical\ event> &\rightarrow <stem>\ <vertical\ event> \\
&\rightarrow <stem>
\end{aligned}
$$

Orientation change from time-prior-to-pitch to pitch-prior-to-time is realized as follows:

$$
\begin{aligned}
<ts\_measure> &\rightarrow <voice\_line>\ <ts\_measure> \\
&\rightarrow <voice\_line> \\
<voice\_line> &\rightarrow <stem>\ <voice\_line> \\
&\rightarrow <stem>
\end{aligned}
$$

Pitch prior to time grammar

$$<stave> \rightarrow \; <sign\_stave> <vl\_stave>$$

$$<sign\_stave> \rightarrow \; \text{beginning-barline} <bl\_stave> \; | \; <bl\_stave>$$

$$<bl\_stave> \rightarrow \; <clef> <cl\_stave> \; | \; <cl\_stave>$$

$$<cl\_stave> \rightarrow <key\_signature> <ks\_stave> \; | \; <ks\_stave>$$

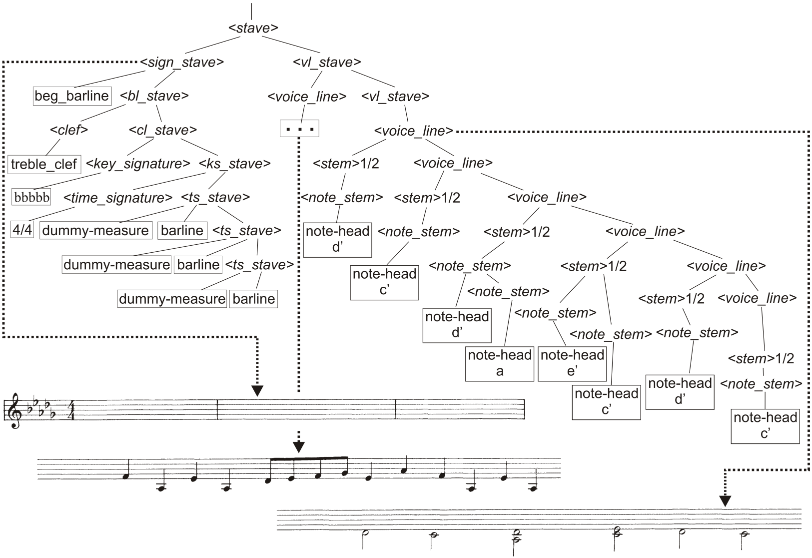
$$<ks\_stave> \rightarrow \; <time\_signature> <ts\_stave> \; | \; <ts\_stave>$$

$$\begin{aligned} <ts\_stave> &\rightarrow \; \text{dummy-measure barline} <ts\_stave> \\ &\rightarrow \; \text{dummy-measure barline} \\ &\rightarrow \; \text{dummy-measure} \end{aligned}$$

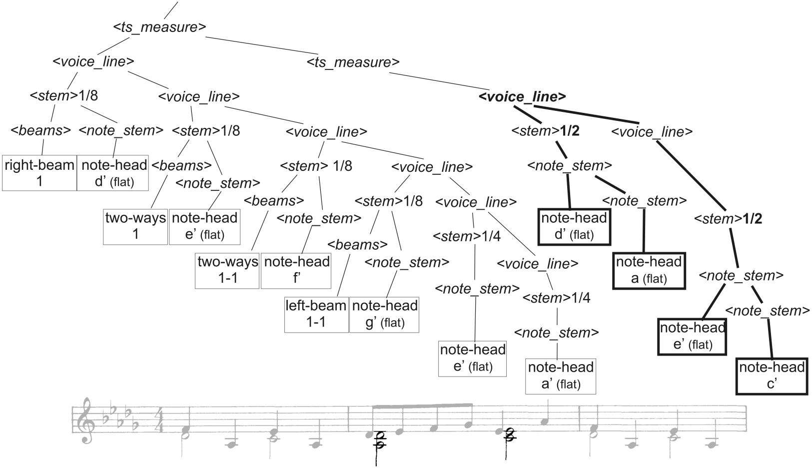
$$<vl\_stave> \rightarrow <voice\_line> <vl\_stave> \; | \; <voice\_line>$$

$$\begin{aligned} <voice\_line> &\rightarrow \; <stem> <voice\_line> \\ &\rightarrow \; <stem> \end{aligned}$$

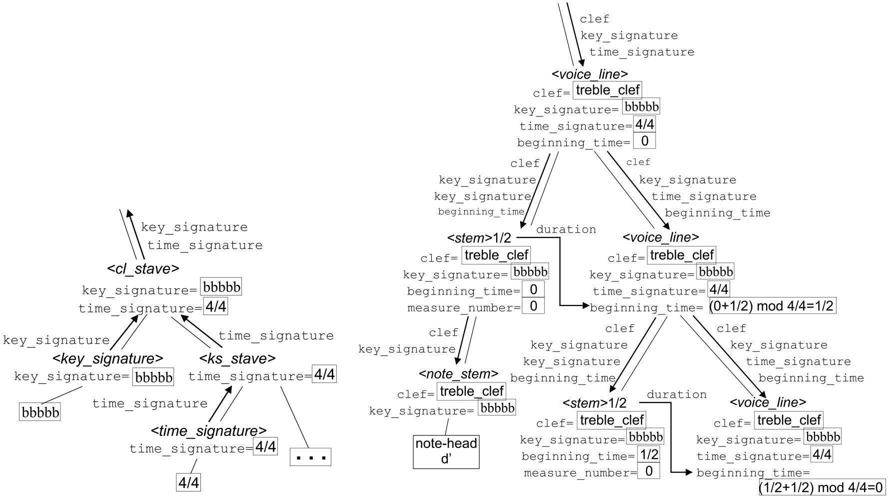
Syntactic Alternatives



Pitch prior to time



Synchronization - attribute grammars



Automatic Data Understanding

the Tool for Intelligent Man-Machine Communication

Władysław Homenda

Faculty of Mathematics and Information Science, Warsaw University of Technology
Faculty of Economics and Informatics in Vilnius, University of Bialystok

SEMAPRO, Barcelona, November 12th, 2017