Dealing Efficiently with Ontology-Enhanced Linked Data for Multimedia

HiDeSt 2015, Dresden

Oliver Gries, Ralf Möller, Anahita Nafissi, Maurice Rosenfeld, Kamil Sokolski, Sebastian Wandelt

Universität zu Lübeck
Kill-App for Stream Data Processing

• Automatic video analysis processes
  – produce sets of multimodal descriptions in a data-driven fashion

• Still, analysis results must be extended
  – to better capture what humans associate with videos and videos shots (→ video interpretation)
  – Asynchronous heterogeneous streams of triples (audio, video, video OCR), plus huge domain kb
  – Foci of attention (centered around video shots)
From a window to a bag of triples
From bags of time-tagged triples to sequences
Prerequisite: Representation of Multimedia Content
Interpretation

Person(p) to Speech(s): hasMediaDecomposition

Person(p) to Speech(s): hasLogicalDecomposition

Person(p) to Speech(s): hasSegmentLocator

Person(p) to Speech(s): hasEnd

Person(p) to Speech(s): hasStart

Person(p) to Speech(s): hasInterviewer

Person(p) to Speech(s): hasContent

Person(p) to Speech(s): depicts

Person(p) to Speech(s): personToSpeech(p,s)
Scalability

- Streams -> windows -> sequences of states
- Each state processed w.r.t. large Abox
- Desideratum: Subdivision of large Aboxes into smaller, meaningful parts
- Solution: Partitioning technique
  - Scalability
  - Window and sequence building
$\mathcal{O}$-Separability

- A role is $\mathcal{O}$-separable with respect to an ontology $\mathcal{O} = (\mathcal{T}, \mathcal{A})$, if semantics for decision problems are preserved when splitting up $R(a, b)$.
Example of $\mathcal{O}$-Separability (1)

\[
\mathcal{T} = \{\text{AudioContent} \sqsubseteq \forall \text{hasMediaDecomposition}.\text{AudioSegment}, \\
\text{AudioSegment} \sqsubseteq \text{MultimediaSegment}\} \\
\mathcal{A}_1 = \{\text{AudioContent}(ac), \text{AudioSegment}(as1) \\
\text{hasMediaDecomposition}(ac, as1)\}
\]
Example of $\mathcal{O}$-Separability (2)

$\mathcal{T} = \{\text{AudioContent} \sqsubseteq \forall \text{hasMediaDecomposition}.\text{AudioSegment}, \text{AudioSegment} \sqsubseteq \text{MultimediaSegment}\}$

$\mathcal{A}_2 = \{\text{AudioContent}(ac), \text{MultimediaSegment}(as1), \text{hasMediaDecomposition}(ac, as1)\}$
**O-Separability**

Diagram showing relationships between various entities and their attributes such as 'hasLogicalDecomposition', 'hasMediaDecomposition', and 'hasSegmentLocator'. The diagram includes nodes labeled with terms like 'MD(md)', 'VC(vc)', 'AC(ac)', 'VS(vs2)', 'VS(vs1)', 'AS(as1)', 'AS(as2)', 'VL(vl2)', 'VL(vl1)', 'AL(al1)', 'AL(al2)', and connections indicating 'depicts', 'hasSegmentLocator', and 'personToSpeech(p,s)'.
CASAM: Online Processing
Offline Parallel Processing of Windows
Conclusions and Outlook

• Processing of media segments with STARQL (\(\rightarrow\) fusion of streams)
• Automatic generation of islands in the context of multimedia interpretation (\(\rightarrow\) scalability, window building)
• Interpretation/abduction as part of (\(\rightarrow\) sequence building)
• Representation of discourse
• Islands not only for speed, but also for efficient management of resources (forgetting)
Acknowledgments

• Funded by
  – European Commission
    (contract FP7-217061 CASAM, done)
  – Deutsche Forschungsgemeinschaft
    (contract MO-801-1 PRESINT, done)
  – European Commision
    (contract FP7-318338 Optique, ongoing)