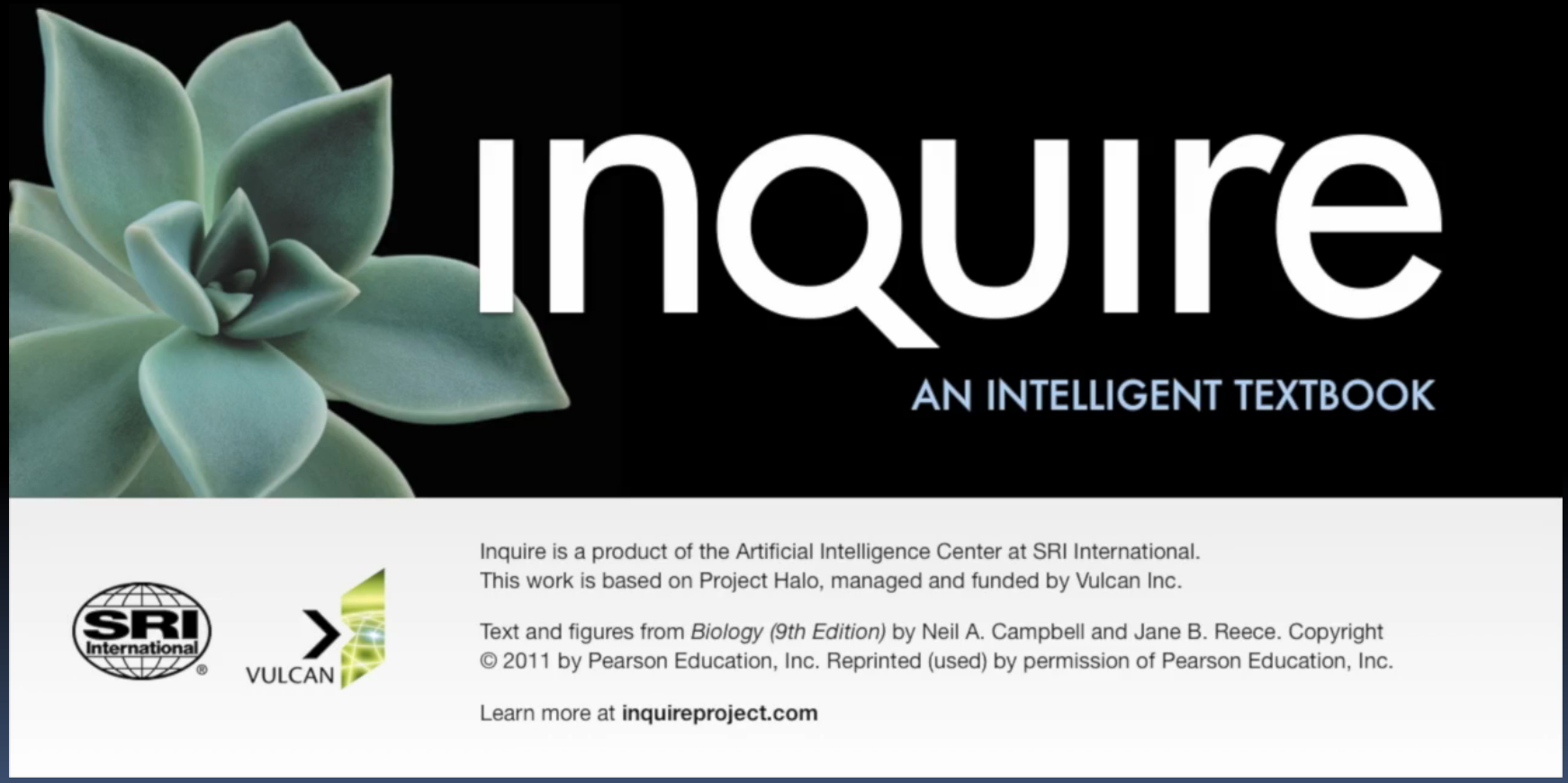
Ralf Möller, Institut für Informationssysteme, Universität zu Lübeck

Logical Foundations for Interpreting Media Data as Streams of Data Descriptions

New Ways of Interacting with Media
New Ways of Teaching?



With permission by Vinay Chaudhri, AI Program Director, Stanford Research Institute

Assumptions and Research Goals

- Way of interacting with media also relevant for professional environments
  - E.g., for tasks in hospitals
- Need to make production of interactive media much less expensive
- Ensure authors have fun producing apps involving media

Generating Symbolical Semantic Content Descriptions for Multimedia Documents

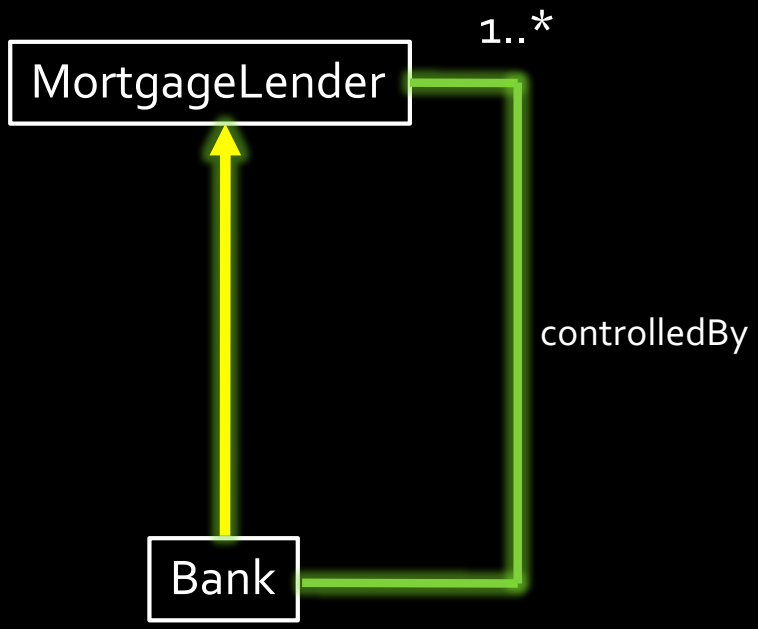
- Establish relations to external resource such as, e.g., Google Knowledge Graph
- Derive relational descriptions of media content
  - Automatic „interpretation“ …
  - … on different layers of abstraction
- Ontologies and inference problems
  - Deduction (Find implicit descriptions)
  - Abduction (Explain „observations“)

Query answering wrt ontologies

Example

controlledBy(BLB, HSH)

Bank(HRE)



Anfrage

{(X)|controlledBy(X,Y)}

?- X=BLB X=HRE, X=HSH

MortgageLender
HRE
HSH
BLB

Bank
HRE
HSH

controlledBy	
BLB	HSH
HRE	??
HSH	??

[Racer 1998 - ...][TONES 05-08]

Symbolic representation of interpretation knowledge: First-order style / deduction

- $\forall$ y, z : Jumper(y), touches(y, z), Pole(z)
  → $\exists$ x : PoleVault(x),
  PV_InStartPhase(x),
  hasPart(x, y),
  hasPart(x, z)

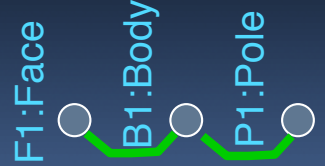
- Hard to realize…
  - No control over first-order prover

Symbolic representation of interpretation knowledge: Datalog style / abduction

$$
\begin{aligned}
touches(Y,Z) \quad &\leftarrow \quad Pole\_Vault(X),\\
&\qquad PV\_InStartPhase(X),\\
&\qquad hasPart(X,Y), Jumper(Y),\\
&\qquad hasPart(X,Z), Pole(Z).\\
near(Y,Z) \quad &\leftarrow \quad Pole\_Vault(X),\\
&\qquad PV\_InEndStartPhase(X),\\
&\qquad hasPart(X,Y), Horizontal\_Bar(Y),\\
&\qquad hasPart(X,Z), Jumper(Z).\\
near(Y,Z) \quad &\leftarrow \quad High\_Jump(X),\\
&\qquad HJ\_InJumpPhase(X),\\
&\qquad hasPart(X,Y), Horizontal\_Bar(Y),\\
&\qquad hasPart(X,Z), Jumper(Z).\\
&\ldots
\end{aligned}
$$

Starting interpretation
Spatial association

Yelena Isinbayeva of Russia on her way to victory (Getty Images)

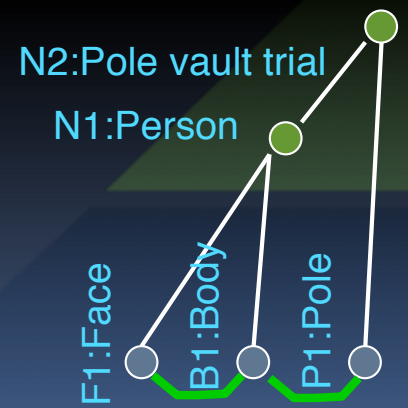


Semantic interpretation
Abduction

Yelena Isinbayeva of Russia on her way to victory (Getty Images)

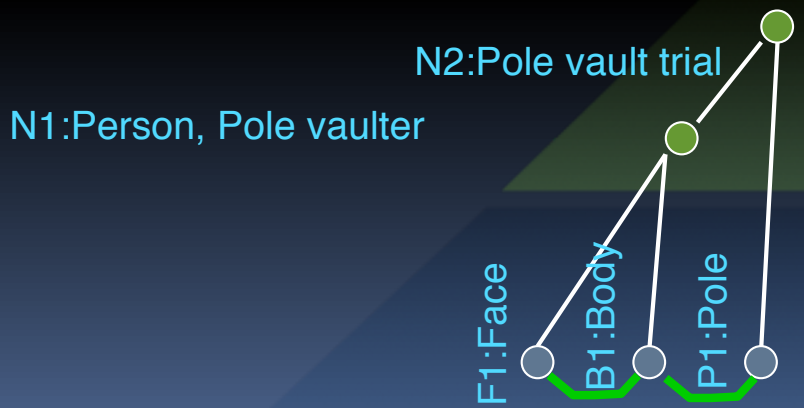
Semantic interpretation
Abduction

Yelena Isinbayeva of Russia on her way to victory (Getty Images)



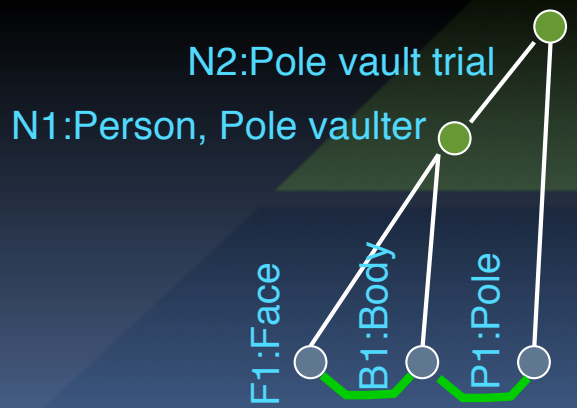
Semantic interpretation
Abduction Deduction

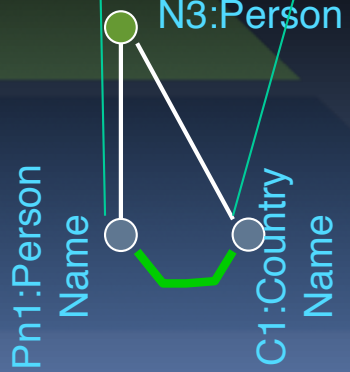
Yelena Isinbayeva of Russia on her way to victory (Getty Images)



Multimedia interpretation
Abduction

Yelena Isinbayeva of Russia on her way to victory (Getty Images)





Multimedia interpretation
Abduction Deduction

Yelena Isinbayeva of Russia on her way to victory (Getty Images)

=

N2:Pole vault trial

N1:Person, Pole vaulter

N3:Person, Pole vaulter

F1:Face

B1:Body

P1:Pole

Pn1:Person Name

C1:Country Name

Multimedia interpretation
Use case: cf. video

Conclusion too unspecific?
Use body atoms as guards!

$$
\begin{aligned}
touches(Y,Z) \quad &\leftarrow \quad Pole\_Vault(X),\\
&\qquad PV\_InStartPhase(X),\\
&\qquad hasPart(X,Y), Jumper(Y),\\
&\qquad hasPart(X,Z), Pole(Z).\\
near(Y,Z) \quad &\leftarrow \quad Pole\_Vault(X),\\
&\qquad PV\_InEndStartPhase(X),\\
&\qquad hasPart(X,Y), Horizontal\_Bar(Y),\\
&\qquad hasPart(X,Z), Jumper(Z).\\
near(Y,Z) \quad &\leftarrow \quad High\_Jump(X),\\
&\qquad HJ\_InJumpPhase(X),\\
&\qquad hasPart(X,Y), Horizontal\_Bar(Y),\\
&\qquad hasPart(X,Z), Jumper(Z).\\
&\ldots
\end{aligned}
$$

Abductive query answering

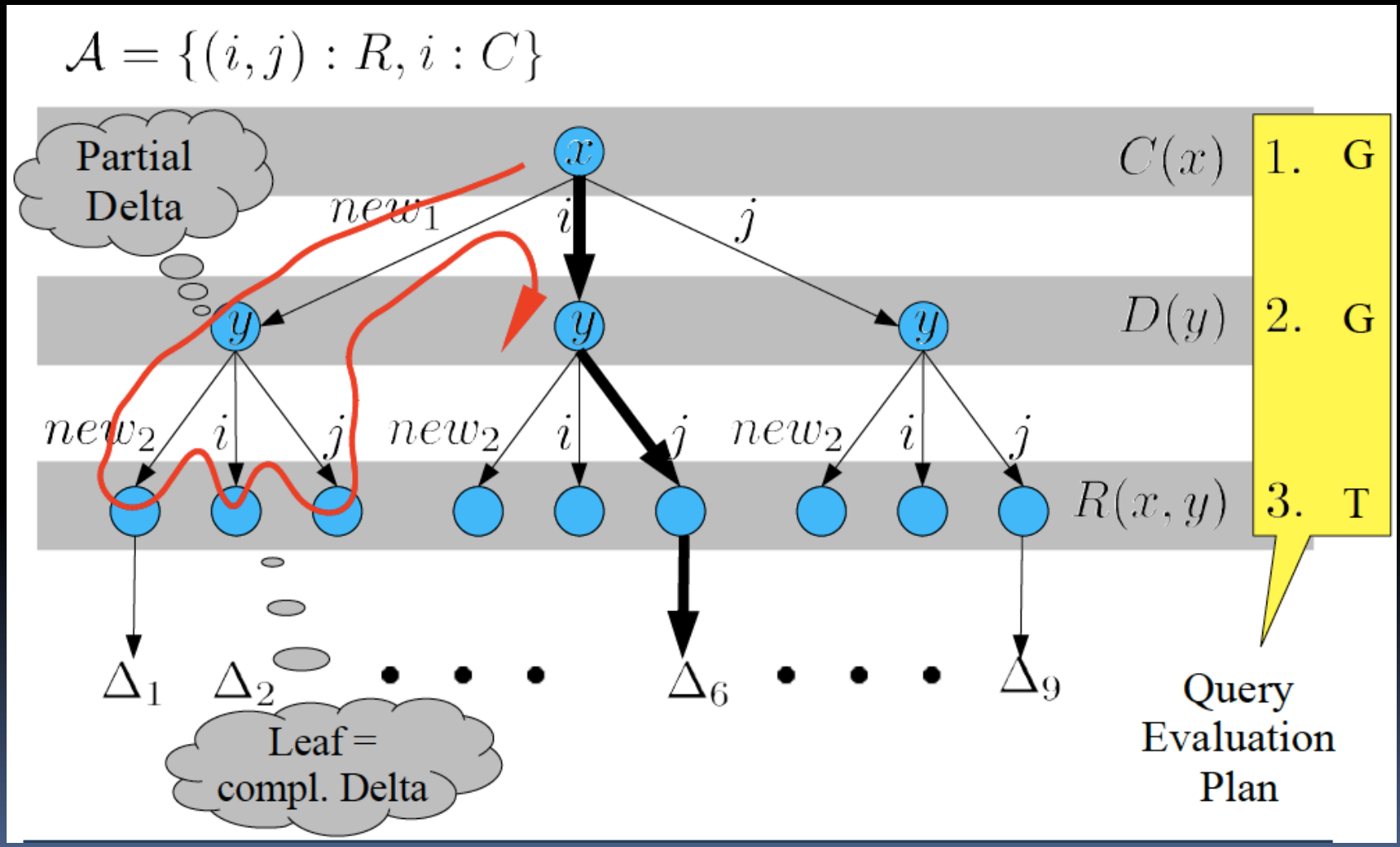
- Simple example
  - Query: $ans() \leftarrow C(x), D(y), R(x, y)$
  - Abox: $\{(i, j) : R, i : C\}$
  - **Preferred** solution (optimal, according to score defined below)

    $x \leftarrow i, y \rightarrow j :$
    $\Delta = \{j : D\}$
  - **Other** solution (plus 7 more, $3^2 = 9$), e.g.

    $x \leftarrow new1, y \leftarrow new_2 :$
    $\Delta = \{new_1 : C, new_2 : D, (new_1, new_2) : R\}$
- Exponential number of solutions has to be computed to find „the best“
  - **optimization idea:** early dynamic cutoff of search space based on score evaluation on partially computed explanations (deltas)

[Diss Wessel: nRQL]

Depth-first abductive query evaluation



[Diss Wessel: nRQL]

Score for comparing solutions

Very simple:

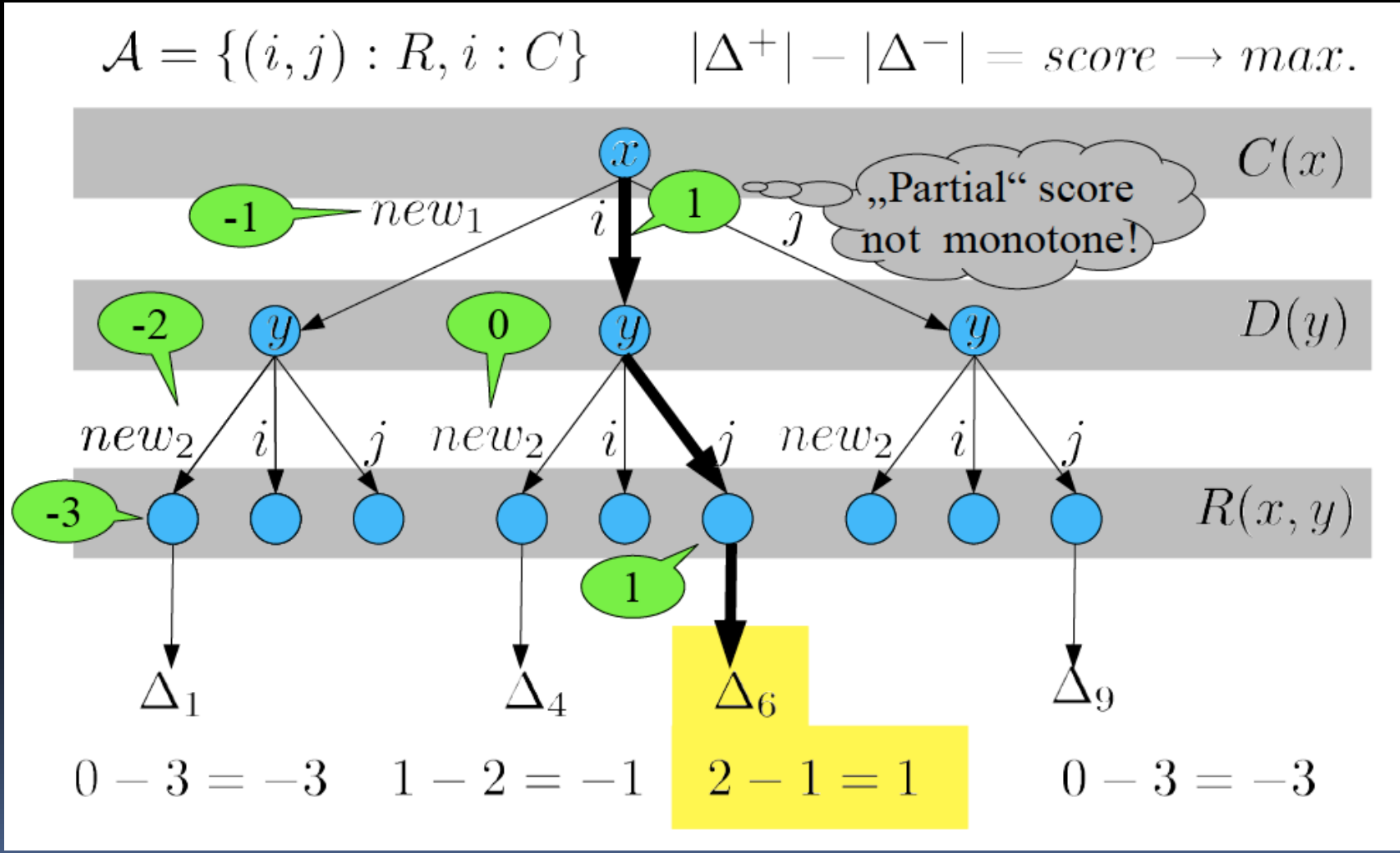
entailed Assertions minus hypothesized Assertions

$$\text{score}(\Delta) =_{def} |\Delta^+| - |\Delta^-| \rightarrow \text{maximize}$$

$$\Delta = \Delta^+ \cup \Delta^- \text{ (entailed, hypothesized)}$$

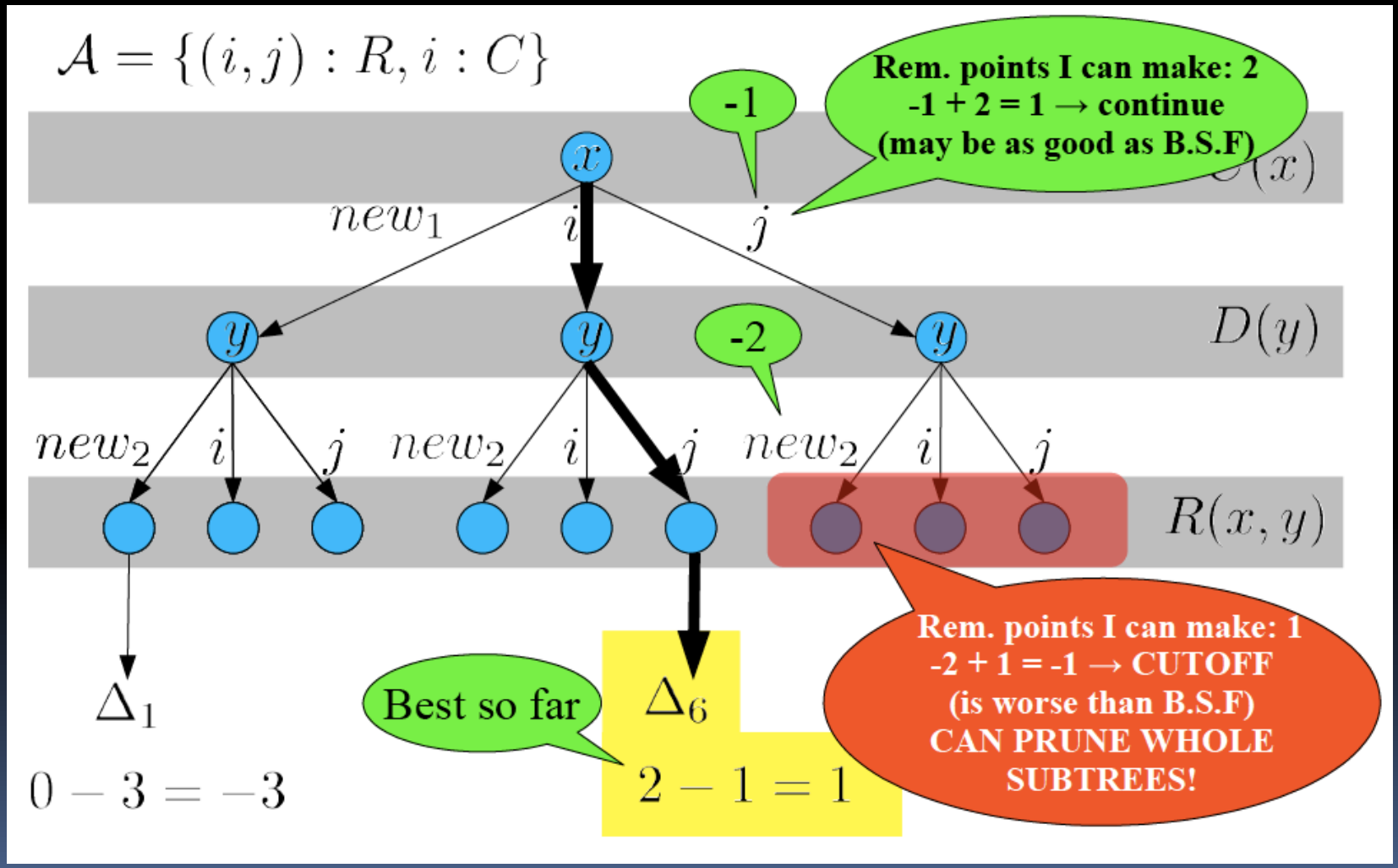
[Diss Wessel: nRQL]

Illustration of partial scores



[Diss Wessel: nRQL]

Score-based cutoff



[Diss Wessel: nRQL]

More formally

$n = |\Delta^+| + |\Delta^-|$ (n const. for each rule body)

$\text{score}(\Delta) =_{def} |\Delta^+| - |\Delta^-| \rightarrow$ maximize (not monotone)

$n + \text{score}(\Delta) = 2|\Delta^+|$

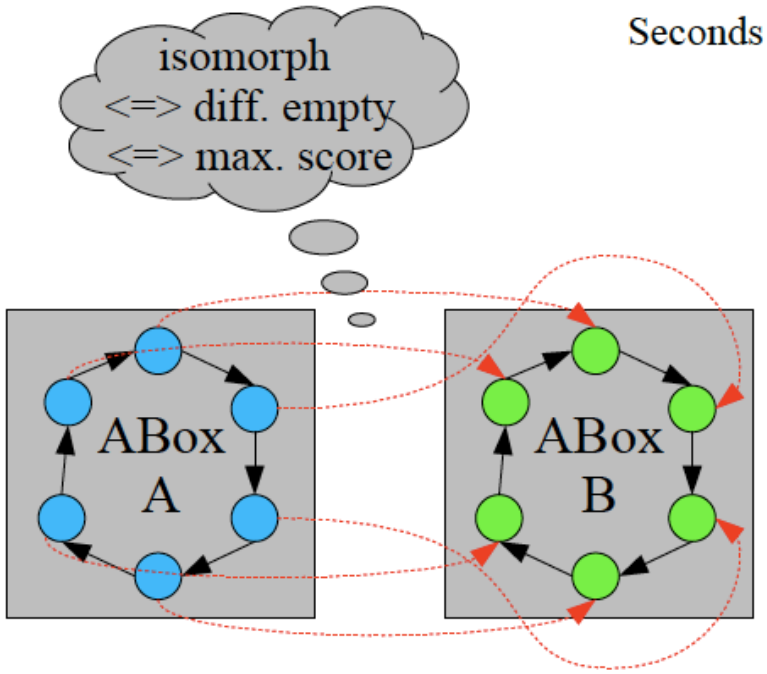
$\text{score}(\Delta) = 2|\Delta^+| - n \rightarrow$ maximize (and monotone!)

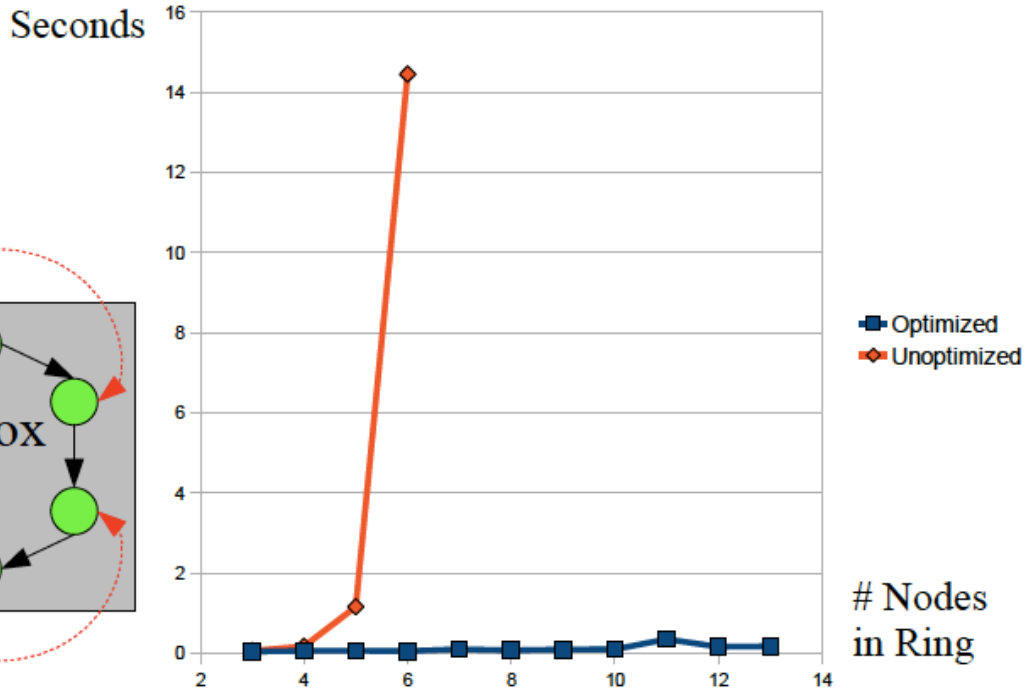
- Let $\Delta_p \subseteq \Delta, m_p = n - |\Delta_p|$ (remaining conjuncts)
  - If $\text{score}(\Delta_p) + (n - |\Delta_p|) < \text{score}(\Delta_{best\_so\_far})$
    $\text{score}(\Delta_{best\_so\_far}) - \text{score}(\Delta_p) > (n - |\Delta_p|)$
    reject Δ_p

[Diss Wessel: nRQL]

How effective is this?

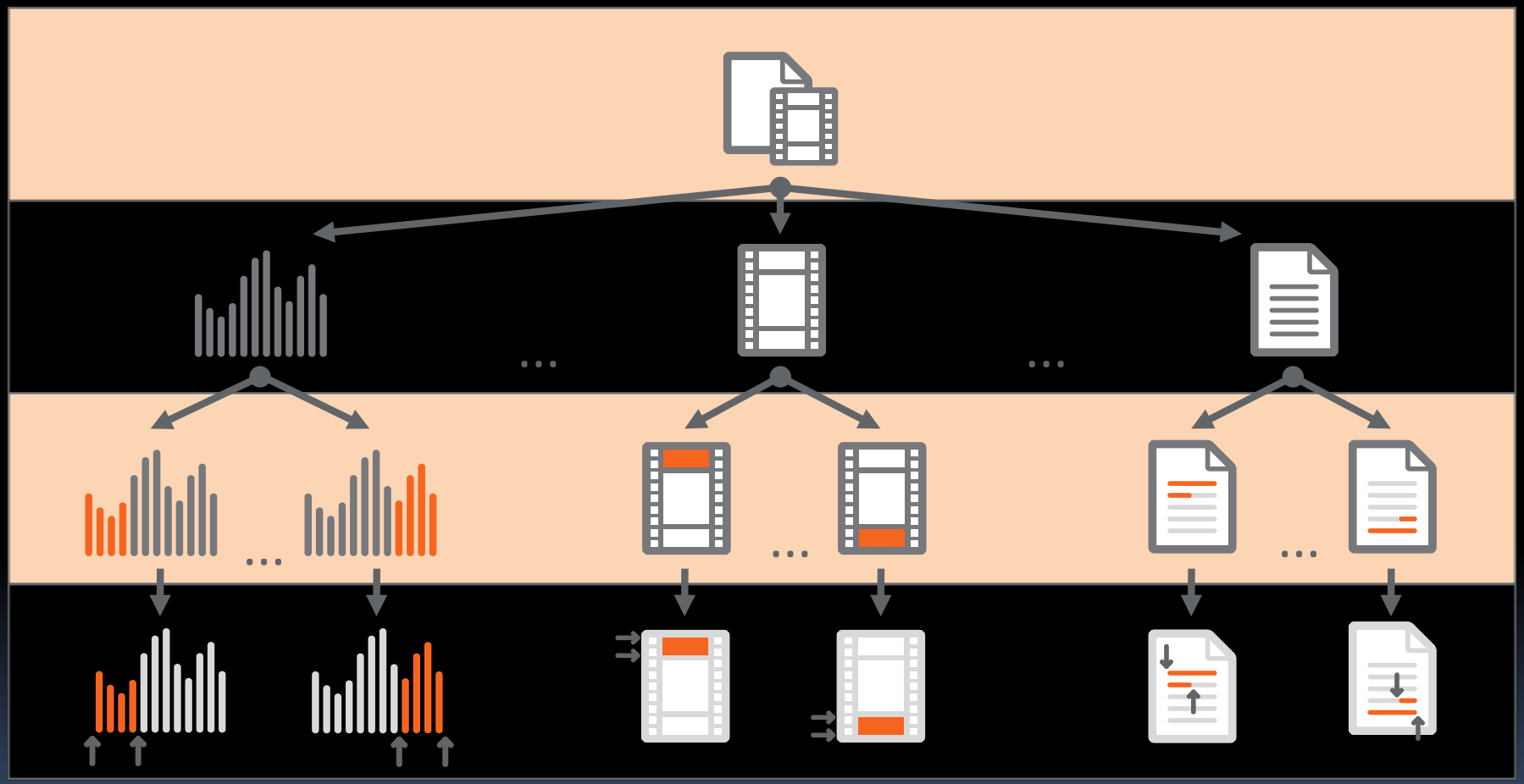
- Synthetic benchmark: finding graph isomorphisms (n nodes)
- Problem reductions:
  Graph Isomorphism → ABox Difference → Abduction





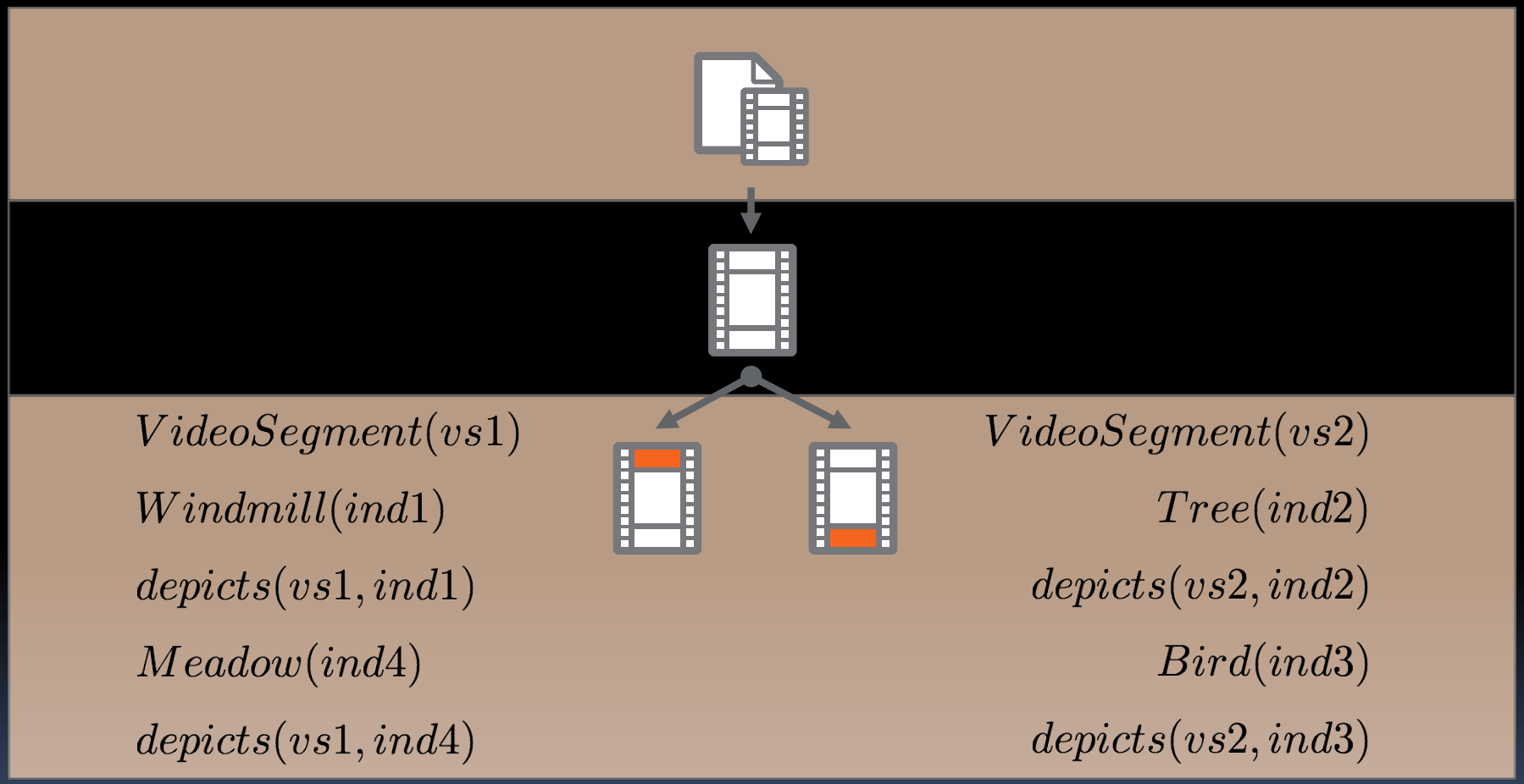
[Diss Wessel: nRQL]

Mulimedia interpretation: Temporal association



Stream-oriented processing (open-world stream)

Mulimedia interpretation: Temporal association



Mulimedia interpretation: Temporal association