Koninklijke Bibliotheek
National library of the Netherlands

vrije Universiteit amsterdam

Using quantitative aspects of alignment generation for argumentation on mappings

Antoine Isaac, Cassia Trojahn,
Shenghui Wang, Paulo Quaresma

Vrije Universteit Amsterdam, University of Evora

Ontology Matching Workshop 2008

Agenda

- Background
- The problem: combining alignments together
- Argumentation frameworks
- Experiments & results
  - Different framework instantiations

Background

- STITCH
  - **S**eman**T**ic **I**nteroperability **T**o access **C**ultural **H**eritage

- Aim: creating alignments between vocabularies from Cultural Heritage

- Semi-formal thesauri
  - As represented in SKOS

Alignment combination problem

- Many alignment techniques and tools available
  - Lexical, structural, etc...
  - #participants in all OAEI campaigns??
- How to select appropriate mappers for a given case?
- Some will perform better than others for this cases
  - Depending on how well the technique fits the vocs. or the application scenario at hand
  - Some perform better for specific parts of the vocabularies to match
- How to combine results of several mappers?
  - Formally distinct but related to selection

Related research

- Recommending mappers
  - Based on profiling (Mochol, 06)
    - Characterizing alignment cases and benchmarking mappers
  - Using sample evaluation/bootstrapping for the case at hand
    - To rank mappers (Tan, 07)
    - To learn composition strategies: weights, thresholds (Ehrig, 05)
  - Problem: all-or-nothing selection

- Filtering individual mappings from alignments
  - Detecting logically inconsistent mappings (Stuckenschmidt, 04)
  - Requires ontologies and mappings with rich formal semantics
  - And we have big vocabularies

What we have still: quantitative aspects of alignment results

- Strength/confidence value
  - The trustfulness of a mapping
  - `(book, publication, exactMatch, 0.7)`
- Consensus
  - The more mappers agree on a given mapping, the more likely it is to be true
  - Cf. OAEI 2007 Food track (van Hage)

Agenda

- Background
- The problem: combining alignments together
- Argumentation frameworks
- Experiments & results
  - Different framework instantiations

A possible option for combining mappers: argumentation

- Not select mappers: just let them agree on what is good
- Focus on individual mappings and contradictions between mappers about them

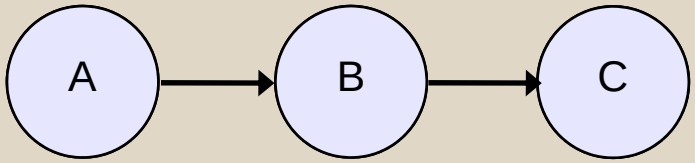
- Already explored for alignment (Laera, 07)
  - With formalized ontologies
- Allows for preferences & strength

- Research question: can we deploy argumentation for cases of informal ontologies, using quantitative aspects of results?

Argumentation framework (Classical)

(Dung, 1995)

- AF = (AR, attacks)
  - *AR*: arguments
  - *attacks*: binary relation over arguments

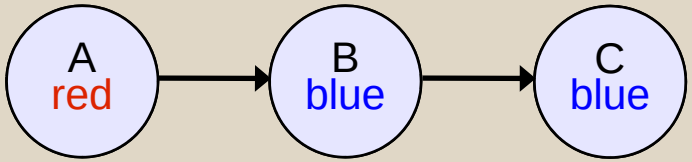


- *Acceptability* of arguments (here, A, C)
  - A is not attacked
  - C is attacked but its attacker is also attacked

Audience-specific AF (VAF)

(Bench-Capon, 2003)

- VAF_{aud}
  - Possible *values*
  - Each argument has a value
  - An audience is associated to a preference order over values



Audience 1: red > blue

Audience 2: blue > red

- *Success* of an attack for an audience
  - A→B is successful for Audience 1
  - B's value is not preferred over A's for Audience 1

Audience-specific strength-based AF

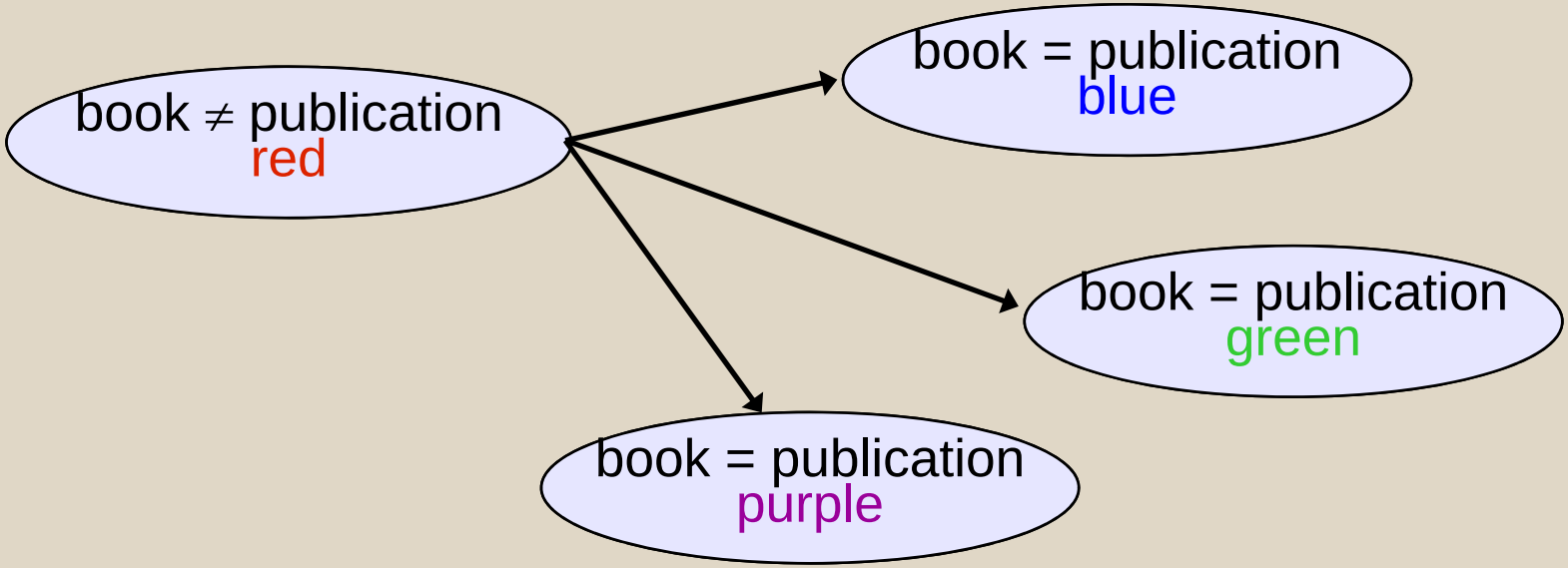
(Trojahn, 2007)

- S-VAF$_{aud}$
  - Each argument has a value and a strength
  - An audience is associated to a preference order over values

- *Success* of an attack (e.g. A→B, B→C)
  - B is stronger than C
  - A is as strong as B but its value is preferred (for Audience 1)

Problem: Consensus?

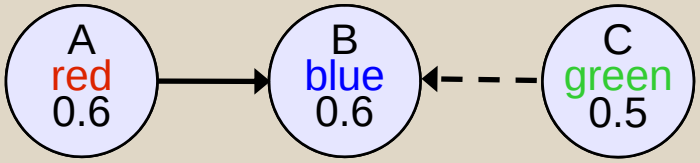
- One single argument can successfully attack any number of arguments
- Even if more of these arguments "support" each other



Introducing voting in argumentation

(us)

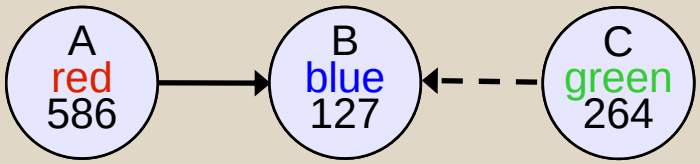
- Sup-VAF$_{aud}$
  - *supports*: (reflexive) binary relation over arguments



- Success of an attack (**plurality voting**)
  - Count supporters of attacker and supporters of attacked
  - Consider preferences when there is a tie
- Raw measure of consensus

Introducing voting in argumentation

- Considering strengths?
  - Problem of scale mismatch
- Comparing "ranks" instead
  - $rank_{map}(A)$= #arguments with strength lower than A's for `map`



- Success of an attack (**borda voting**)
  - Comparing average ranks of attackers and supporters
  - Consider preferences when there is a tie

Agenda

- Background
- The problem: combining alignments together
- Argumentation frameworks
- Experiments & results
  - Different framework instantiations

Experiments - setting

- Context: OAEI library case
  - 2 thesauri
  - exactMatch, broadMatch, relatedMatch
- Mappers
  - OAEI 2007 mappers (Falcon, Silas, DSSim)
  - Home-grown mappers: Instance-based mapper, Dutch lexical mapper, edit-distance mapper
- Evaluation
  - Using mappings to re-annotate books
  - Automatic: books already annotated by 2 vocs

Instantiating frameworks: arguments

- Argument generation

  `A = (c1, c2, s, r, v, h)`

  - c1, c2: mapped concepts
  - s: strength
  - r: type of relationship (e.g. `exactMatch`)
  - v: value representing a mapper (e.g. `instance-based`)
  - h=+ or – : argument is in *favor* or *against* the mapping
    - Allows to define attack and support

Instantiating frameworks: arguments

- State-of-the-art mappers output “positive” mappings
- It's easy to generate positive arguments

  `(book, publication, exactMatch, 0.6)` by instance-based mapper

  → `(book,publication,exactMatch,0.6,instance-based,+)`
- But how to generate negative ones?
  - Related work has exploited formal disjointness
  - But we are in a non-formalized context!

Instantiating frameworks: counter-arguments

2 approaches for attack:

- Negative argument as failure (NAF)
  - create **(c1, c2, r, 1, map, -)** if no **(c1, c2, r, X)** for **map**

  Assumption: *mappers try to give complete results*

- Attack based on disjoint relations (NARD)
  - If there is **(c1, c2, r, s)** for mapper **map**
  - for all mapping relations **r'** that are not **r**
  - generate **(c1, c2, r', s, map, -)**

  Assumption: *different thesaurus links cannot hold between 2 concepts*

Quite bold assumptions, o*bject of experimentation as well!*

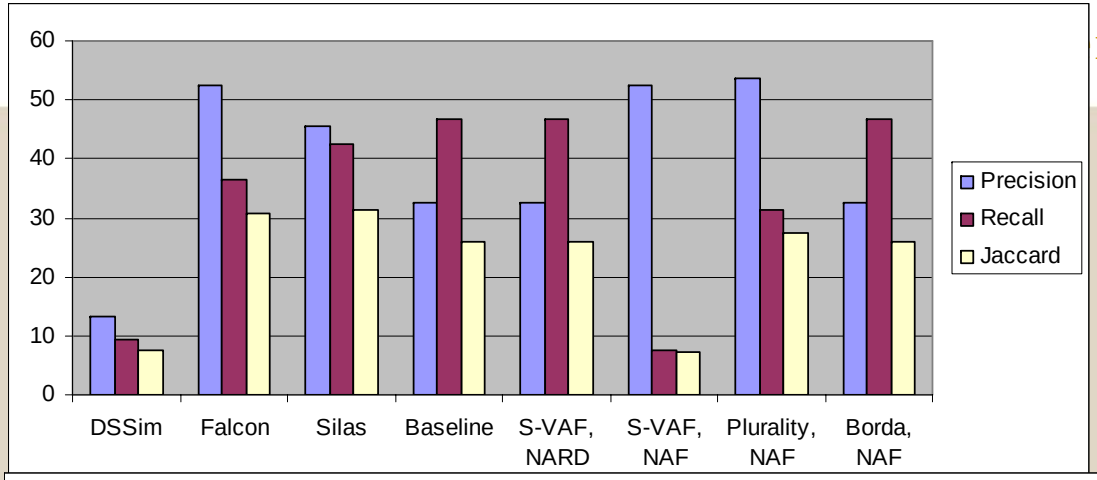
Experiments – combinations and frameworks

- 3 combinations of mappers
  - OAEI, Homegrown mappers, All
- For each combination, different framework tests
  - F2: S-VAF, NARD
  - F1: S-VAF, NAF
  - F3: plurality voting Sup-VAF, NAF
  - F4: Borda voting Sup-VAF, NAF
  - Baseline: simple union of results
- An audience is derived from each mapper

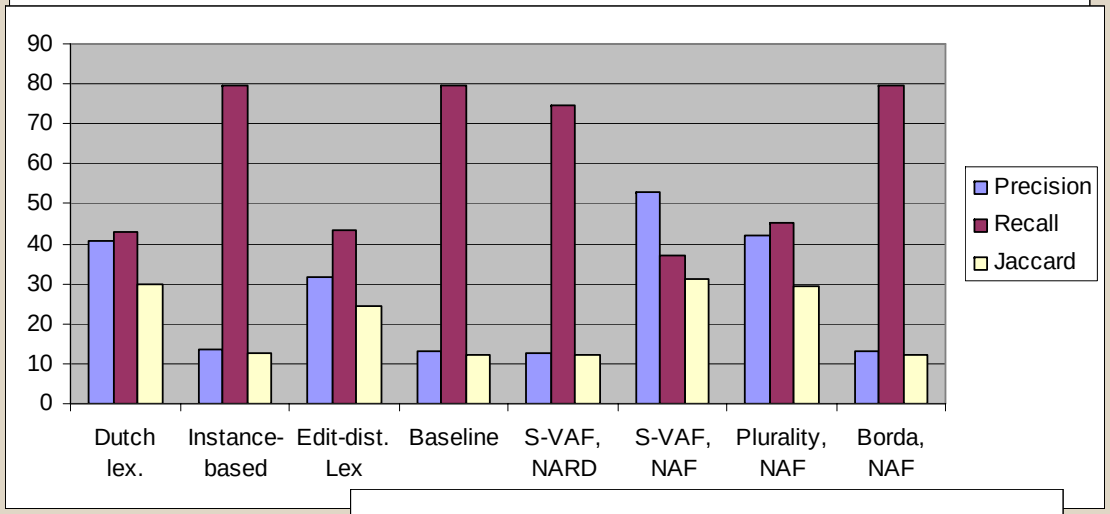
  ≈ Adhoc preference ordering based on individual performances of mappers and self-preference

Results

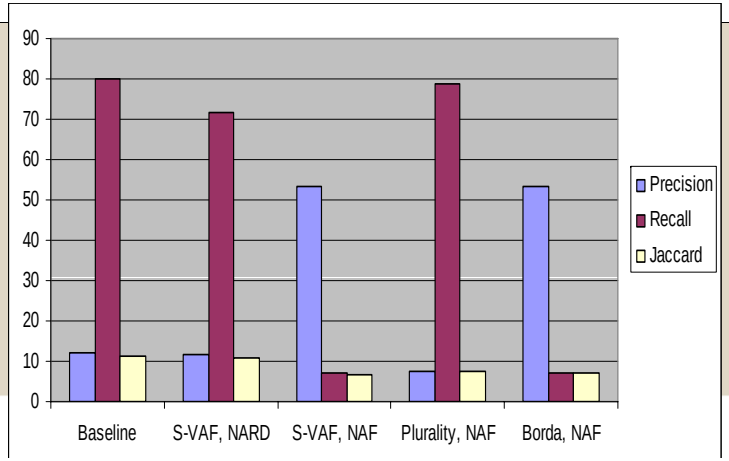
OAEI mappers



Homegrown mappers



All mappers



Discussion

- S-VAF: inconclusive
  - Dependent on negative argument strategy and/or mappers
  - Confirms the problems of comparing strengths across different mappers?
  - NAF amounts most of the time to intersection
    - Due to argumentation setting (objectively acceptable arguments)
- Plurality voting: ≈ OK
- Borda: our implementation does not differ much from S-VAF – NAF

Discussion

- It is possible to enhance on baseline union
  - Gaining on P while not harming too much R
  - Interesting when mappers give lots of imprecise results
- Comparison with best individual: more inconclusive
  - Only F3 consistently enhances P (at the cost of R)
  - Reminder: if we assume that we don't know in advance which one is the best, it is interesting to have comparable results
- Great dependence on mappers involved
  - NARD has not generated lots of attacks (esp. for OAEI)

Future work

- More insight on the process
  - Which proportion of attacked mappings?
  - And successful attacks?
- More « semantic » attack & defense relations
  - Using thesaurus information (hierarchy)
- Experimentation with other aggregation methods
  - But there's already a lot of options available...

Thanks

- Questions?