Experience in Aligning Anatomical Ontologies

Olivier Bodenreider
Lister Hill National Center for Biomedical Communications
Bethesda, Maryland - USA
Acknowledgments

◆ Songmao Zhang

- National Library of Medicine, USA
- Academy of Mathematics and System Sciences, Chinese Academy of Sciences, P.R. China
Govard Bidloo
Gérard de Lairesse
1690

Outline

◆ Direct alignment
  ● Concepts (FMA-GALEN)
  ● Relationships
  ● Cross-species (Human-Mouse)

◆ Indirect alignment through a reference ontology
Aligning Anatomical Concepts using Lexical and Structural Methods
Introduction

◆ Different representations of one domain
  ● Formalism
  ● Structure
  ● Domain coverage

◆ Various approaches for comparing representations
  ● Merging
  ● Transformation
  ● Alignment
Introduction

◆ Objective

- Aligning two representations of anatomy
  - Foundational Model of Anatomy
  - GALEN common reference model
- Beyond lexical comparison
- Investigating reasoning capabilities

◆ Related work

- P. Mork, et al., Challenges in precisely aligning models of human anatomy using generic schema matching

Materials
Two representations of anatomy

**FMA**
- Foundational Model of Anatomy
- University of Washington, 1994
- Conceptualization of the physical objects and spaces that constitute the human body

**GALEN common reference model**
- Generalized Architecture for Languages, Encyclopaedias and Nomenclatures in medicine
- University of Manchester, 1991
- Development of a compositional and generative formal system for modeling all and only sensible medical concepts
## FMA and GALEN

<table>
<thead>
<tr>
<th></th>
<th>FMA</th>
<th>GALEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underlying data model</td>
<td>Frame-based structure</td>
<td>Description logic</td>
</tr>
<tr>
<td></td>
<td>Protégé-2000</td>
<td>GRAIL</td>
</tr>
<tr>
<td>Domain coverage</td>
<td>Anatomy</td>
<td>Medicine</td>
</tr>
<tr>
<td>Concepts</td>
<td>Pre-coordinated</td>
<td>Post-coordinated</td>
</tr>
<tr>
<td></td>
<td>59,422</td>
<td>25,192</td>
</tr>
<tr>
<td>Synonyms</td>
<td>28,686</td>
<td>No</td>
</tr>
<tr>
<td>Anonymous concepts</td>
<td>No</td>
<td>3,176</td>
</tr>
<tr>
<td>Hierarchical relationships</td>
<td><em>is-a, part-of</em> (6)</td>
<td><em>is-a, part-of</em> (26)</td>
</tr>
<tr>
<td>Multiple inheritance</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Associative relationships</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Methods
Alignment steps

Lexical alignment

- Step 1: Acquiring terms
- Step 2: Identifying anchors (i.e., shared concepts) lexically

Structural alignment

- Step 3: Acquiring (explicit and implicit) semantic relations
- Step 4: Identifying anchors structurally
Step 1: Acquiring terms

◆ Extracting concept names

  ● FMA

    ■ Preferred concept names
      - *Uterine tube*

    ■ Synonyms
      - *Oviduct*

  ● GALEN

    ■ Non-anonymous concept names
      - *RoundLigamentOfUterus*

    ▶ Not considered: anonymous concepts
      - *(BodyStructure which < HasDivision Muscle >)*
Step 2: Identifying anchors lexically

◆ Comparing two systems at the term level
  • Exact match
  • Normalization

◆ Preferred concept names and synonyms

◆ Anchors (*i.e.*, shared concepts)
  • FMA: *Fibularis tertius*
    (synonym: *Peroneus tertius*)
  • GALEN: *Peroneus Tertius*
Step 3: Acquiring semantic relations

◆ Semantic relations
  - \(<concept_1, \text{relationship}, \text{concept}_2>\)
  - Hierarchical relationships: is-a and part-of
    - \(<\text{Arm}, \text{part-of}, \text{Proximal segment of upper limb}>\)

◆ Extracting the explicit relations

◆ Acquiring implicit knowledge
  - Complementing missing inverse relations
  - Augmenting relations embedded in concept names
  - Inferring relations from a combination of relations
Implicit knowledge Complementing

- Inverse relationships
  - *is-a* and *inverse-is-a*
  - *part-of* and *has-part*

- Most *part-of* relations not represented bi-directionally
  - *<Arm, has-part, Humerus>*

- Complementing the inverse relations
  - *<Humerus, part-of, Arm>*
Implicit knowledge  Augmenting

◆ Reification of *part-of* relationships

\[ <X, \text{is-	extit{a}}, \text{Part of } Y> \]

\[ <X, \text{part-of}, Y> \]

◆ Augmenting reified *part-of* relations

- Reified: \(<\text{Cardiac chamber, is-	extit{a}, Subdivision of heart}>\>
- No explicit (direct or indirect) *part-of* relationships between \textit{Cardiac chamber} and \textit{Heart} in FMA
- Augmented: \(<\text{Cardiac chamber, part-of, Heart}>\>
Generating new inter-concept relationships by applying inference rules
Step 4: Identifying anchors structurally

- Structural similarity: common relations among anchors

**FMA**

- Muscle of leg
  - Plasma membrane
    - has-part
    - has-part
  - Muscle cell
  - Leg

**GALEN**

- Muscle of Leg
  - Fibrous Tissue
    - has-part
  - Myocyte
  - Leg

- Fibularis tertius
  - has-part
  - part-of

- Peroneous Tertius
  - has-part
  - part-of
Conflicts

- Conflicts: semantic incompatibility between anchors
  - Opposite type of links
    - FMA: *Wall of heart has-part* Apex of heart
    - GALEN: *Heart Wall part-of* Apex of Heart
  - Disjoint top-level categories
    - FMA: *Foot is-a* Anatomical structure
      - GALEN: *feet is-a* Unit
Results
Anchors identified by lexical alignment

- 2,353 anchors
  - 4 % of FMA concepts
  - 9 % of GALEN concepts
Semantic relations acquired

<table>
<thead>
<tr>
<th>Types of relations</th>
<th>FMA</th>
<th>GALEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicitly represented</td>
<td>238,135</td>
<td>214,403</td>
</tr>
<tr>
<td>Complemented</td>
<td>104,754</td>
<td>107,689</td>
</tr>
<tr>
<td>Augmented</td>
<td>315,860</td>
<td>27,274</td>
</tr>
<tr>
<td>Inferred</td>
<td>5,172,668</td>
<td>1,661,824</td>
</tr>
<tr>
<td>Total</td>
<td>5,831,417</td>
<td>2,011,190</td>
</tr>
<tr>
<td>Anchors identified by structural alignment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><img src="image_url" alt="Image" /></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No evidence</strong></td>
<td><strong>Positive evidence</strong></td>
<td><strong>Negative evidence</strong></td>
</tr>
<tr>
<td><img src="image_url" alt="Image" /></td>
<td><img src="image_url" alt="Image" /></td>
<td><img src="image_url" alt="Image" /></td>
</tr>
</tbody>
</table>
Discussion
Explicit vs. implicit knowledge

- More positive structural evidence found for anchors
- Augmentation accounted for 74% of 523 anchors acquiring positive evidence
- More conflicting relations found for anchors
References


References


Comparing Associative Relationships among Equivalent Concepts across Ontologies
Lobular organ

is-a

Pancreas

Liver

Lung

Exocrine pancreas

has-part

Neck of Pancreas

bounded by

Surface of Pancreas

arterial supply

Dorsal Pancreatic Artery

Left lung

Right lung
Introduction

* Few ontology merging / aligning tools deal with the issue of comparing associative relationships

* Our ontology aligning project
  - Two representations of anatomy
    - Foundational Model of Anatomy (FMA)
    - GALEN Common Reference Model
  - Aligning hierarchical relationships manually
  - Aligning concepts based on both lexical and hierarchical similarity
Introduction

◆ Objective: to identify equivalent expressions for associative relationships across ontologies

◆ Assumptions
  ● Correspondence
    ■ between two relationships
    ■ between one relationship and a combination of relationships
  ● Types of match
    ■ one-to-one
    ■ one-to-many
    ■ no match
  ● Frequency of the correspondence
Materials
Two representations of anatomy

◆ **FMA**
  - Foundational Model of Anatomy
  - University of Washington, 1994
  - Conceptualization of the physical objects and spaces that constitute the human body

◆ **GALEN common reference model**
  - Generalized Architecture for Languages, Encyclopaedias and Nomenclatures in medicine
  - University of Manchester, 1991
  - Development of a compositional and generative formal system for modeling all and only sensible medical concepts
## FMA and GALEN

<table>
<thead>
<tr>
<th></th>
<th>FMA</th>
<th>GALEN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Underlying data model</strong></td>
<td>Frame-based structure</td>
<td>Description logic</td>
</tr>
<tr>
<td></td>
<td>Protégé-2000</td>
<td>GRAIL</td>
</tr>
<tr>
<td><strong>Domain coverage</strong></td>
<td>Anatomy</td>
<td>Medicine</td>
</tr>
<tr>
<td><strong>Concepts</strong></td>
<td>Pre-coordinated</td>
<td>Post-coordinated</td>
</tr>
<tr>
<td></td>
<td>66,879</td>
<td>52,006</td>
</tr>
<tr>
<td><strong>Hierarchical relationships</strong></td>
<td><em>is-a, part-of</em> (8)</td>
<td><em>is-a, part-of</em> (40)</td>
</tr>
<tr>
<td>Inverses</td>
<td><em>inverse-is-a, has-part</em></td>
<td><em>inverse-is-a, has-part</em></td>
</tr>
<tr>
<td><strong>Associative relationships</strong></td>
<td>59</td>
<td>562</td>
</tr>
<tr>
<td></td>
<td>Some have inverses</td>
<td>Every one has inverse</td>
</tr>
<tr>
<td></td>
<td><em>branch of / branch</em></td>
<td><em>isBranchOf / hasBranch</em></td>
</tr>
<tr>
<td></td>
<td>input from</td>
<td><em>isServedBy / serves</em></td>
</tr>
</tbody>
</table>
Methods
Comparing associative relationships

- NOT based on lexical similarity
- Based on previously identified equivalent concept pairs between FMA and GALEN
  - Share both lexical and hierarchical similarity
    - FMA: Pancreas has-part Exocrine pancreas
    - GALEN: Pancreas has-part ExocrinePancreas
  - Anchors (i.e., equivalent concepts across ontologies)
    - 2,604 pairs
    - 4% of FMA concepts and 5% of GALEN concepts
Step 1 Acquiring associative relations

- Associative relations: \( \text{concept}_1 \text{ relationship } \text{concept}_2 \)

- Extracting the explicit relations
  - \text{Kidney is Served By AutonomicNerveOfAbdomen}

- Complementing the missing inverse relations
  - \text{AutonomicNerveOfAbdomen serves Kidney}

- Augmenting relations embedded in concept names
  - Explicit: \( X \text{ branch-of } Y \text{ is a } \text{Branch of musculocutaneous nerve} \)
  - Augmented: \( X \text{ is a } \text{Branch of Y}\text{ branch of Musculocutaneous nerve} \)
Step 2 Identifying relationship patterns

- Search for inter-anchor path pairs

```
FMA

Pancreas

isServedBy

Caudal Pancreatic Artery

isBranchOf

Inferior Pancreatic Artery

isBranchOf

Dorsal Pancreatic Artery

GALEN
```
Step 2 Identifying relationship patterns

- Create relationship patterns from path pairs

**FMA**

- Pancreas
  - Dorsal Pancreatic Artery
  - Caudal Pancreatic Artery
  - Inferior Pancreatic Artery

**GALEN**

- Pancreas
  - Caudal Pancreatic Artery
  - Inferior Pancreatic Artery
  - Dorsal Pancreatic Artery

*arterial supply*
Step 2 Identifying relationship patterns

◆ Create relationship patterns from path pairs
  ● FMA: *arterial supply*
  ● GALEN: *isServedBy* – *isBranchOf*

◆ Direct and indirect relationship patterns

◆ Frequency of relationship pattern
  ■ *Number of path pairs sharing the pattern*  
    
  ■ Most frequent vs. accidental relationship patterns
Results
### Associative relations acquired

<table>
<thead>
<tr>
<th>Associative relations</th>
<th>FMA</th>
<th>GALEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicit</td>
<td>18,688</td>
<td>288,732</td>
</tr>
<tr>
<td>Complemented</td>
<td>1,057</td>
<td>249,938</td>
</tr>
<tr>
<td>Augmented</td>
<td>1,838</td>
<td>108</td>
</tr>
<tr>
<td>Total</td>
<td>21,583</td>
<td>538,778</td>
</tr>
<tr>
<td>Between anchors</td>
<td>847</td>
<td>6,922</td>
</tr>
</tbody>
</table>
Path pairs and relationship patterns identified

- 4,070 inter-anchor path pairs
- 350 relationship patterns (47 direct and 303 indirect)

Frequency distribution of relationship patterns
## Examples of relationship patterns

<table>
<thead>
<tr>
<th>FMA</th>
<th>GALEN</th>
<th>Frequency (N = 4,070)</th>
</tr>
</thead>
<tbody>
<tr>
<td>part-of</td>
<td>isBranchOf</td>
<td>518 13%</td>
</tr>
<tr>
<td>branch of</td>
<td>isBranchOf</td>
<td>310  8%</td>
</tr>
<tr>
<td>tributary of</td>
<td>isBranchOf</td>
<td>104  3%</td>
</tr>
<tr>
<td>member of</td>
<td>is-a</td>
<td>42  1%</td>
</tr>
<tr>
<td>nerve supply</td>
<td>part-of – isServedBy</td>
<td>16  0.4%</td>
</tr>
<tr>
<td>part-of – contained in</td>
<td>isNonPartitivelyContainedIn</td>
<td>10  0.25%</td>
</tr>
<tr>
<td>contained in</td>
<td>boundsSpace – inverse-isA</td>
<td>2  0.05%</td>
</tr>
</tbody>
</table>
## Multiple matches

<table>
<thead>
<tr>
<th>FMA</th>
<th>GALEN</th>
<th>Frequency (N = 74)</th>
</tr>
</thead>
<tbody>
<tr>
<td>arterial supply</td>
<td><code>isServedBy</code></td>
<td>18 24%</td>
</tr>
<tr>
<td></td>
<td><code>isServedBy – isBranchOf</code></td>
<td>34 46%</td>
</tr>
<tr>
<td></td>
<td><code>isServedBy – part-of</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>isServedBy – is-a</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>isServedBy – inverse-isa</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other combinations</td>
<td>22 30%</td>
</tr>
</tbody>
</table>
Discussion
## Analysis of relationship patterns

<table>
<thead>
<tr>
<th>Types of patterns</th>
<th>Number of patterns (N = 350)</th>
<th>Examples</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Associative</td>
<td>14 4%</td>
<td>F: <em>tributary of</em> G: <em>isBranchOf</em></td>
<td>Equivalent associative relationships</td>
</tr>
<tr>
<td>Associative and Hierarchical</td>
<td>33 9%</td>
<td>F: <em>bounded by</em> G: <em>has-part</em></td>
<td>Not semantic equivalence</td>
</tr>
<tr>
<td>Associative and Combination</td>
<td>303 87%</td>
<td>F: <em>arterial supply</em> G: <em>isServedBy – is-a</em></td>
<td>Different levels of granularity or modeling choices</td>
</tr>
</tbody>
</table>
Semantic vs. lexical correspondence

- **Semantically and lexically similar**
  - 3 cases
  - \{FMA: \textit{branch of}, GALEN: \textit{isBranchOf}\}

- **Semantically similar but lexically different**
  - 11 cases
  - \{FMA: \textit{arterial supply}, GALEN: \textit{isServedBy}\}

- **Semantically different but lexically similar**
  - 4 cases
  - \{FMA: \textit{bounded by}, GALEN: \textit{isSpaceBoundedBy}\}
Limitations and future work

- Associative relationships do not exist in paths between anchors are not matched
  - 56% of FMA (e.g., *fascicular architecture*)
  - 84% of GALEN (e.g., *isPositionedDistalTo*)

- Anchors used for identifying equivalent relationships have not been fully validated

- Take advantage of the equivalent relationships to discover more equivalent concepts
References

  
Aligning Mouse and Human Anatomies
Introduction

◆ Anatomy is central to the biomedical domain
◆ Comparing functional information about genes across model organisms requires aligned anatomies
◆ Objective: to align two ontologies of anatomy
  - Mouse anatomy
    *Adult Mouse Anatomical Dictionary*
  - Human anatomy
    *Anatomy subset of NCI Thesaurus*
◆ Contribution to the caBIG project
Materials
Adult Mouse Anatomical Dictionary (MA)

- Structured controlled vocabulary
- 2,404 concepts each identified by one name
  - *Head/neck, Adrenal artery*
- 259 synonyms
  - *Limb* has a synonym *Extremity*
- Directed acyclic graph
- Two relationships: *is-a* and *part-of*
- 38% concepts have no *is-a* relationship
  - *Knee* *part-of* *Hindlimb*
- 4% concepts have more than one *is-a* relationship
  - *Hand phalanx* *is-a* *Phalanx*
    - *is-a Hand digit bone*
NCI Thesaurus (NCI)

- Standard vocabularies for cancer research
- Anatomy class
- Available in Ontology Web Language (OWL)
- 4,410 concepts, each having one preferred name
  - *Abdominal esophagus*
- 2,371 synonyms
  - *Orbit* has a synonym *Eye socket*
- Every concept has at least one *is-a* relationship
- 4% concepts have more than one *is-a* relationship
  - *Radius bone* is-a *Long bone* is-a *Bone of the upper extremity*
- Concepts are connected by a *part-of* relationship
  - *Liver* is physical part of *Gastrointestinal system*
## MA and NCI

<table>
<thead>
<tr>
<th></th>
<th>MA</th>
<th>NCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underlying data model</td>
<td>Directed acyclic graph</td>
<td>Available in OWL</td>
</tr>
<tr>
<td>Domain coverage</td>
<td>Mouse anatomy</td>
<td>Human anatomy related to cancers</td>
</tr>
<tr>
<td>Concepts</td>
<td>2,404</td>
<td>4,410*</td>
</tr>
<tr>
<td>Synonyms</td>
<td>259</td>
<td>2,371</td>
</tr>
<tr>
<td>Hierarchical relationships</td>
<td>is-a, part-of</td>
<td>is-a, part-of</td>
</tr>
<tr>
<td>Inverses</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* Of which some 2000 correspond to entities not included in MA
Methods
Overview

Lexical alignment

Manual alignment

Lexical mapping

Manual mapping

Structural validation

Structural validation

NCI

MA

NCI

MA

Lexical alignment

Manual alignment

Evaluation (manual)
Lexical approach

◆ Comparing two ontologies at the term level
  • Exact match
  • Match after normalization

◆ Preferred names and synonyms are used
  • MA: Forelimb
    NCI: Upper extremity (synonym: Forelimb)

◆ UMLS synonymy is used to identify additional matches
  • MA: Profunda femoris artery
    NCI: Deep femoral artery
Validation by structural similarity

- Uses relations explicitly represented in each system and transitive closures
- Presence of relations to other anchors interpreted as structural evidence
Evaluation

Lexical alignment

Manual alignment

Shared matches

Specific matches (reviewed)

Specific matches (reviewed)
Results
Results

NCI MA

Lexical alignment

Manual alignment

Lexical mapping

Manual mapping

Structural validation

Structural validation

Lexical alignment in common

639 mappings in common
Comparison of the two alignments

- Lexical alignment: 715
- Manual alignment: 781
- Overlap: 639
Mappings identified by both approaches

- **639 mappings identified by both approaches**
  - Most of them supported by structural evidence
    - {MA: *uterine cervix*, NCI: *Cervix Uteri*}
  - Some not supported by structural evidence
    - {MA: *tendon*, NCI: *Tendon*}

---

Diagram:

- MA: **tendon** → **Connective tissue**
  - **Tendon** → **Aponeurosis**
  - **Tendon** → **Musculoskeletal System**

NCI:
Mappings specific to the lexical approach

- 76 mappings specific to the lexical approach
  - Benefited from using UMLS synonyms
  - 61 valid mappings (80%)
    - \{MA: lienal artery, NCI: Splenic Artery\}
  - 15 invalid mappings
    - \{MA: cerebellum lobule I, NCI: Lingula of the Lung\}
Mappings specific to the manual approach

◆ 142 mappings specific to the manual approach
  ● 133 valid mappings (94%)
    ■ {MA: alveolus epithelium, NCI: Alveolar Epithelium}
  ● 9 invalid mappings
    ■ Human errors (coding)
Discussion
Applications of the mapping for biologists

◆ **Important for comparative science**
  
  ● Mouse models of human diseases
  
  e.g., [emice.nci.nih.gov](http://emice.nci.nih.gov)

◆ **Example**

```
<table>
<thead>
<tr>
<th>Human</th>
<th>Mouse</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anatomical site</strong>&lt;sub&gt;H&lt;/sub&gt;</td>
<td><strong>Anatomical site</strong>&lt;sub&gt;M&lt;/sub&gt;</td>
</tr>
<tr>
<td><strong>Disease</strong>&lt;sub&gt;H&lt;/sub&gt;</td>
<td><strong>Disease</strong>&lt;sub&gt;M&lt;/sub&gt;</td>
</tr>
</tbody>
</table>
```
Lessons learned

◆ Curated mapping
  • Only one expert

◆ Lexical approach
  • Large proportion of valid mappings
    ■ Including among mappings not supported by structural evidence (conservative approach)

◆ Manual approach
  • Can be supported by automated validation techniques (structural evidence), used to focus the attention of experts on potential problems
References

Indirect Alignment of Multiple Ontologies of Anatomy: through a Reference Ontology
Approaches to aligning multiple ontologies

Pairwise alignment

Alignment through a reference
**Introduction**

- **Objective:** to investigate the indirect alignment of two anatomical ontologies through a reference ontology

![Diagram showing indirect alignment between two ontologies](image-url)
Introduction

◆ Three ontologies of anatomy:
  ● Adult Mouse Anatomical Dictionary (MA)
  ● Anatomy subset of NCI Thesaurus (NCI)
  ● Foundational Model of Anatomy (FMA)

◆ First attempt to automatically derive mappings among ontologies from their alignments to a reference ontology
Materials
Adult Mouse Anatomical Dictionary

◆ Structured controlled vocabulary
◆ 2,404 concepts each identified by one name
  ● Head/neck, Adrenal artery
◆ 259 synonyms
  ● Limb has a synonym Extremity
◆ Directed acyclic graph
◆ Two relationships: is-a and part-of
◆ 38% concepts have no is-a relationship
  ● Knee part-of Hindlimb
◆ 4% concepts have more than one is-a relationship
  ● Hand phalanx is-a Phalanx
  is-a Hand digit bone
NCI Thesaurus

- Standard vocabularies for cancer research
- Anatomy class
- Available in Ontology Web Language (OWL)
- 4,410 concepts, each having one preferred name
  - Abdominal esophagus
- 2,371 synonyms
  - Orbit has a synonym Eye socket
- Every concept has at least one is-a relationship
- 4% concepts have more than one is-a relationship
  - Radius bone is-a Long bone
    - is-a Bone of the upper extremity
- Concepts are connected by a part-of relationship
  - Liver is physical part of Gastrointestinal system
Foundational Model of Anatomy

- Conceptualize the physical objects and spaces that constitute the human body
- Frame-based structure in Protégé
- 71,202 concepts, each having one preferred name
  - Uterine tube
- 52,713 synonyms
  - Uterine tube has a synonym Oviduct
- Every concept has one and only one is-a relationship
- Seven part-of relationships and their inverses
  - constitutional part of and constitutional part
  - regional part of and regional part
Methods
Three phases

Direct alignment
MA-FMA

Direct alignment
NCI-FMA

Alignment through the FMA

Direct alignment
MA-NCI
Phase 1 Direct alignment

Lexical alignment

- Acquiring terms
- Identifying matches (i.e., shared concepts) lexically

Structural alignment

- Acquiring (explicit and implicit) semantic relations
- Identifying matches structurally
Phase 2 Indirect alignment

- Upper limb
- Forelimb
- Upper extremity

FMA

- Forelimb

MA

NCI

direct

indirect

(direct)
Phase 3 Comparison of two alignments

Direct alignment

Indirect alignment

Specific matches

Shared matches

Specific matches
Results
Indirect alignment

- MA-FMA: 1,353 matches in direct alignment
- NCI-FMA: 2,173 matches in direct alignment
- MA-NCI: 715 matches in direct alignment
- NCI-MA: 703 matches in alignment through the FMA
Comparison of two alignments

Direct alignment

Indirect alignment

715

61

654

703

49
Discussion
Benefits of the indirect alignment

- Why are the 49 matches not identified in the direct alignment?
  - Additional synonyms by the FMA
    - MA: *Integumental system*
    - NCI: *Integumentary system*
  - Additional relations by the FMA
    - MA: *Hip bone*
    - NCI: *Pelvic bone*
    - NCI: *Ischium* (syn: *Hip*)
    - FMA: *Hip Bone*
    - FMA: *Hip*
Benefits of the direct alignment

Why are the 61 matches not identified in the indirect alignment through the FMA?

- Different coverage

**MA:**
- Common iliac artery
- Internal iliac artery
- External iliac artery
- Iliac artery

**FMA:**
- Common iliac artery
- Internal iliac artery
- External iliac artery

**NCI:**
- Common iliac artery
- Internal iliac artery
- External iliac artery
- Iliac artery
Benefits of the direct alignment

- Why are the 14 matches supported in the direct alignment while having no evidence in the indirect alignment through the FMA?
  - Different representation

![Diagram showing the relationships between Blood vessel, Artery, and Vein in MA, FMA, and NCI]

MA:
- Blood vessel
- Artery
- Vein

FMA:
- Blood vessel
- Artery
- General anatomical term
- Vein

NCI:
- Blood vessel
- Artery
- Vein
Alignment through a reference vs. pairwise alignment

◆ Efficiency of alignment through a reference
  • $n(n-1)/2$ pairwise mappings
  • $(n-1)$ mappings to a reference

◆ Feasibility of alignment through a reference
  • Identified 91% of matches in the direct alignment
  • Identified additional matches not discovered by the direct alignment
  • Validated the FMA as a reference ontology
References

Medical Ontology Research

Contact: olivier@nlm.nih.gov
Web: mor.nlm.nih.gov

Olivier Bodenreider
Lister Hill National Center
for Biomedical Communications
Bethesda, Maryland - USA