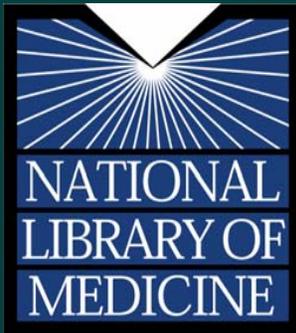


Medical Informatics Training Program
Final Presentation
March 30, 2005

Aligning Representations of Anatomical Knowledge



Songmao Zhang, Ph.D.

Mentor: Olivier Bodenreider, M.D., Ph.D.

Lister Hill National Center
for Biomedical Communications

Bethesda, Maryland - USA

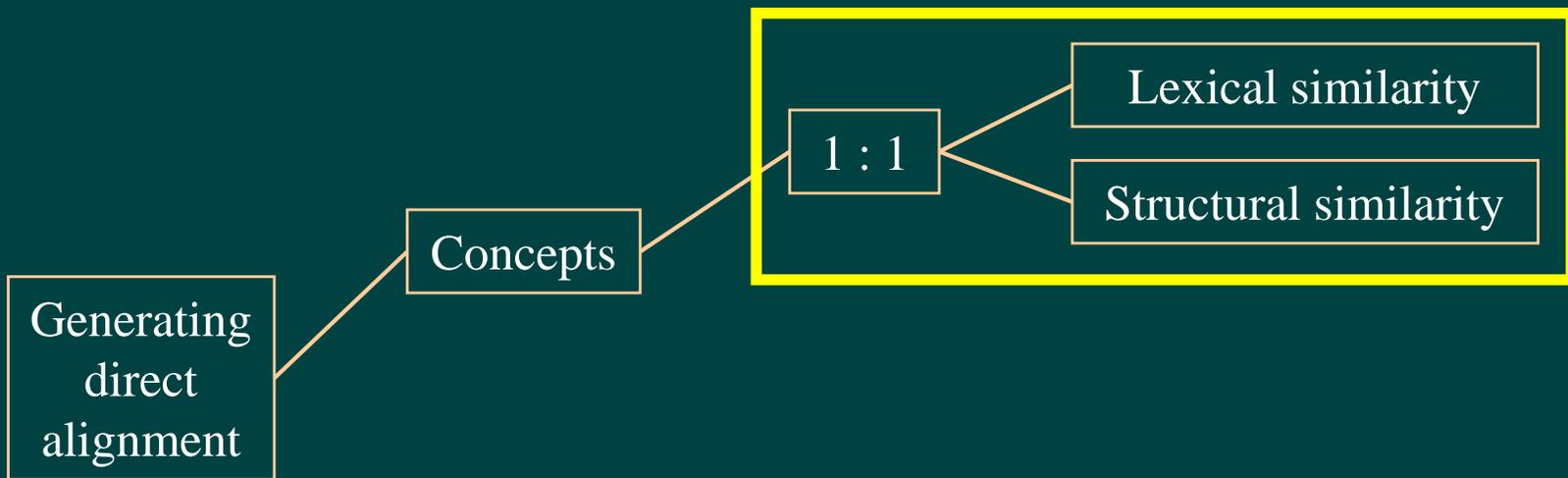
General framework

- ◆ Comparison and evaluation of structures and reasoning potentials in different medical knowledge representation systems

- ◆ Case study:
 - Foundational Model of Anatomy (FMA)
 - Frame-based structure
 - GALEN common reference model (GALEN)
 - Description logic



*An overview
of our alignment project*



1:1 concept match

◆ Lexical alignment

FMA: *Fibularis tertius* (synonym: *Peroneous tertius*)

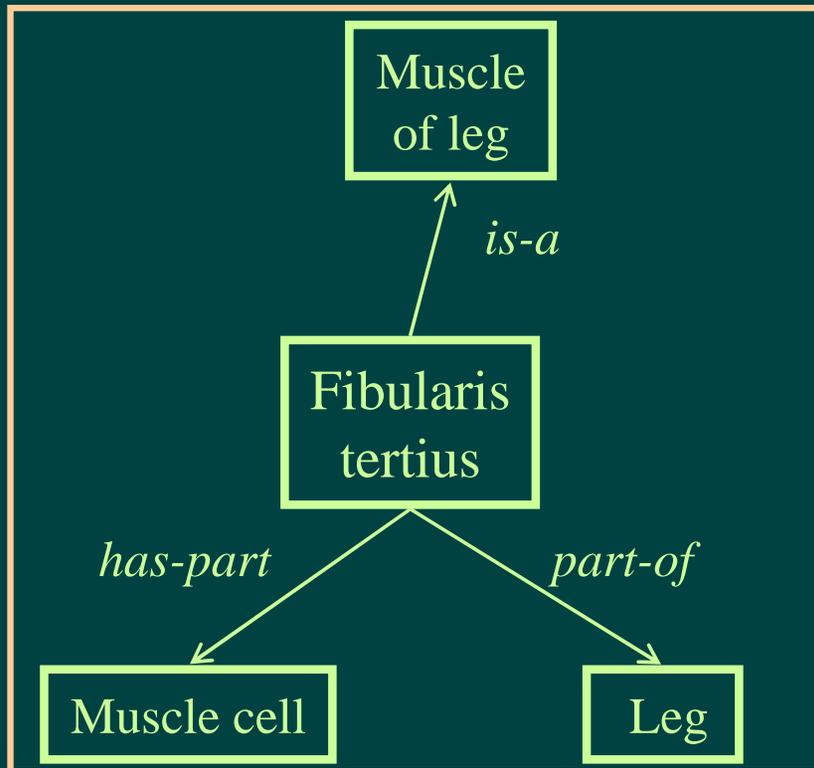
GALEN: *Peroneus Tertius*



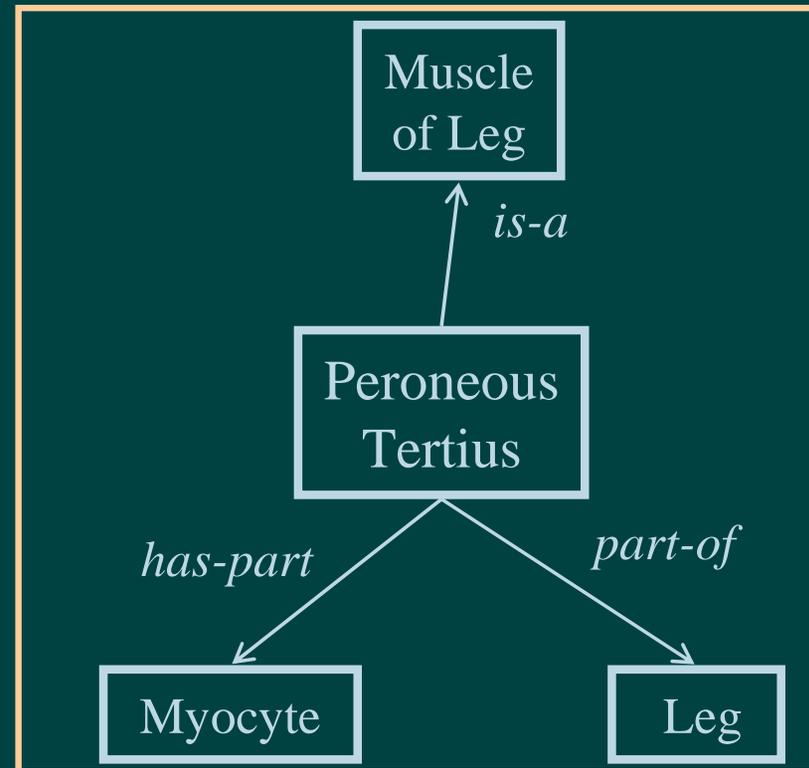
1:1 concept match

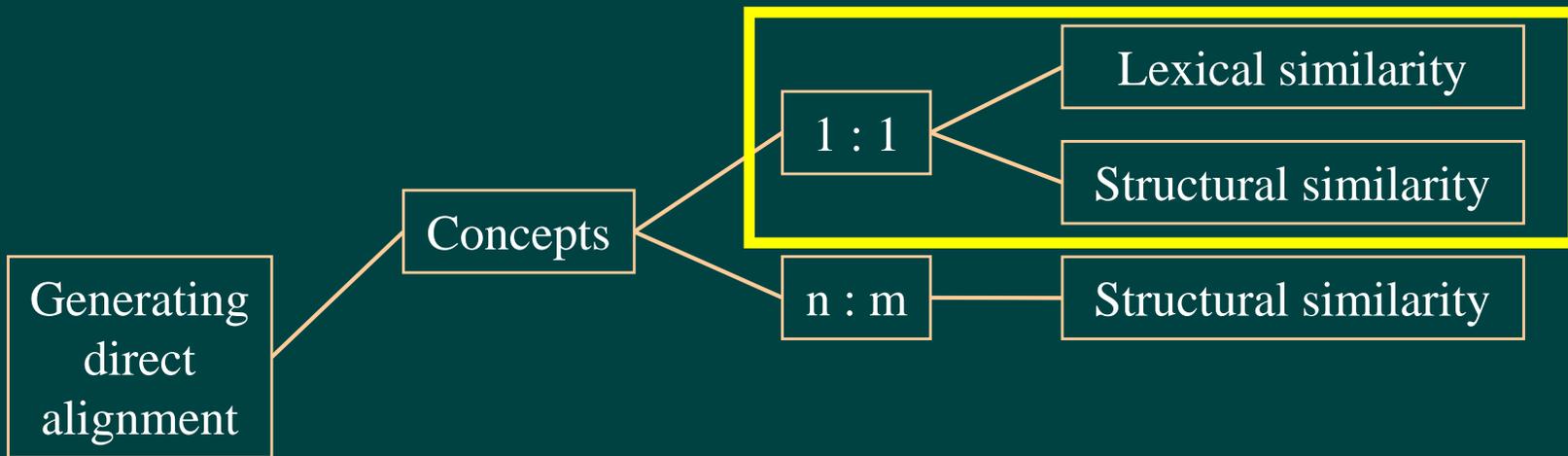
- ◆ Structural alignment: shared hierarchical relationships among concept

FMA



GALEN

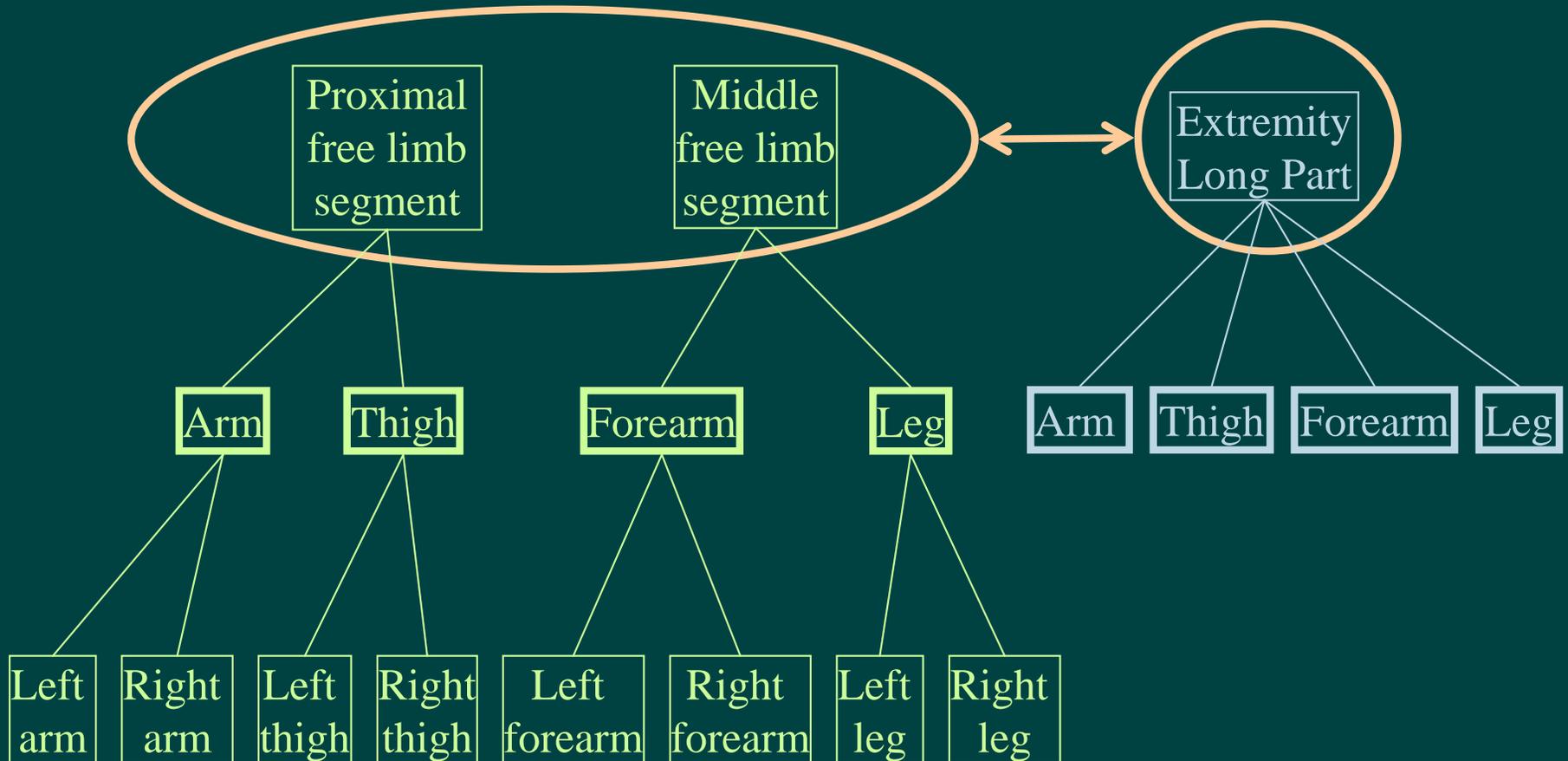


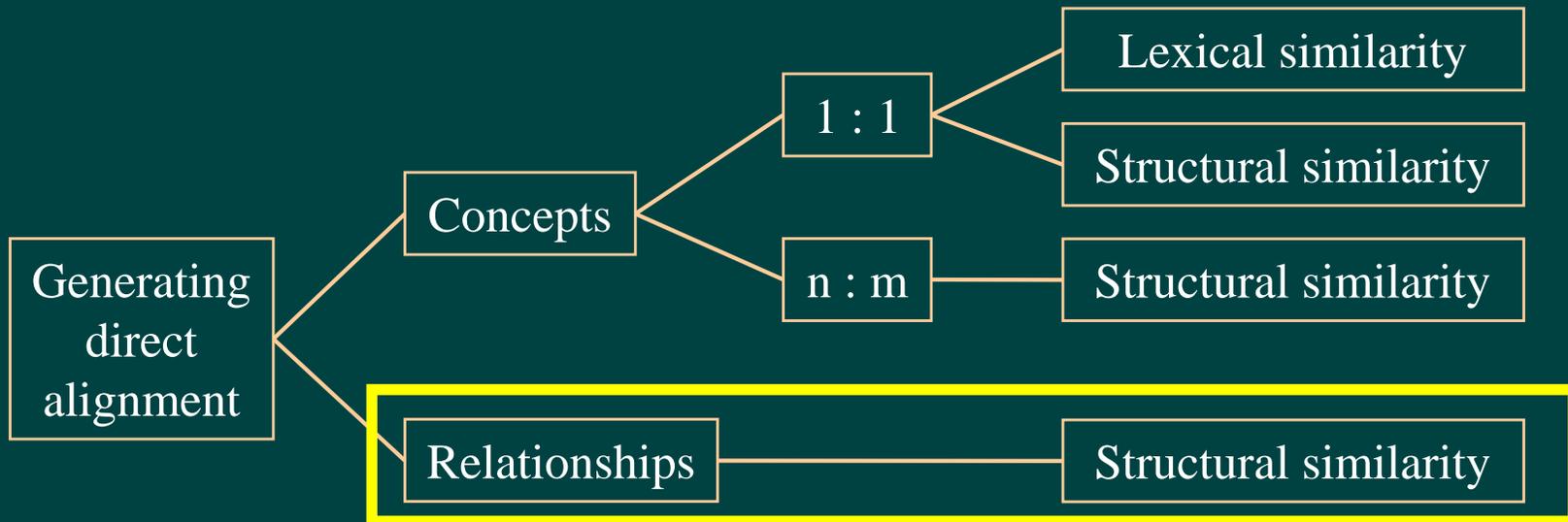


n:m concept match

FMA

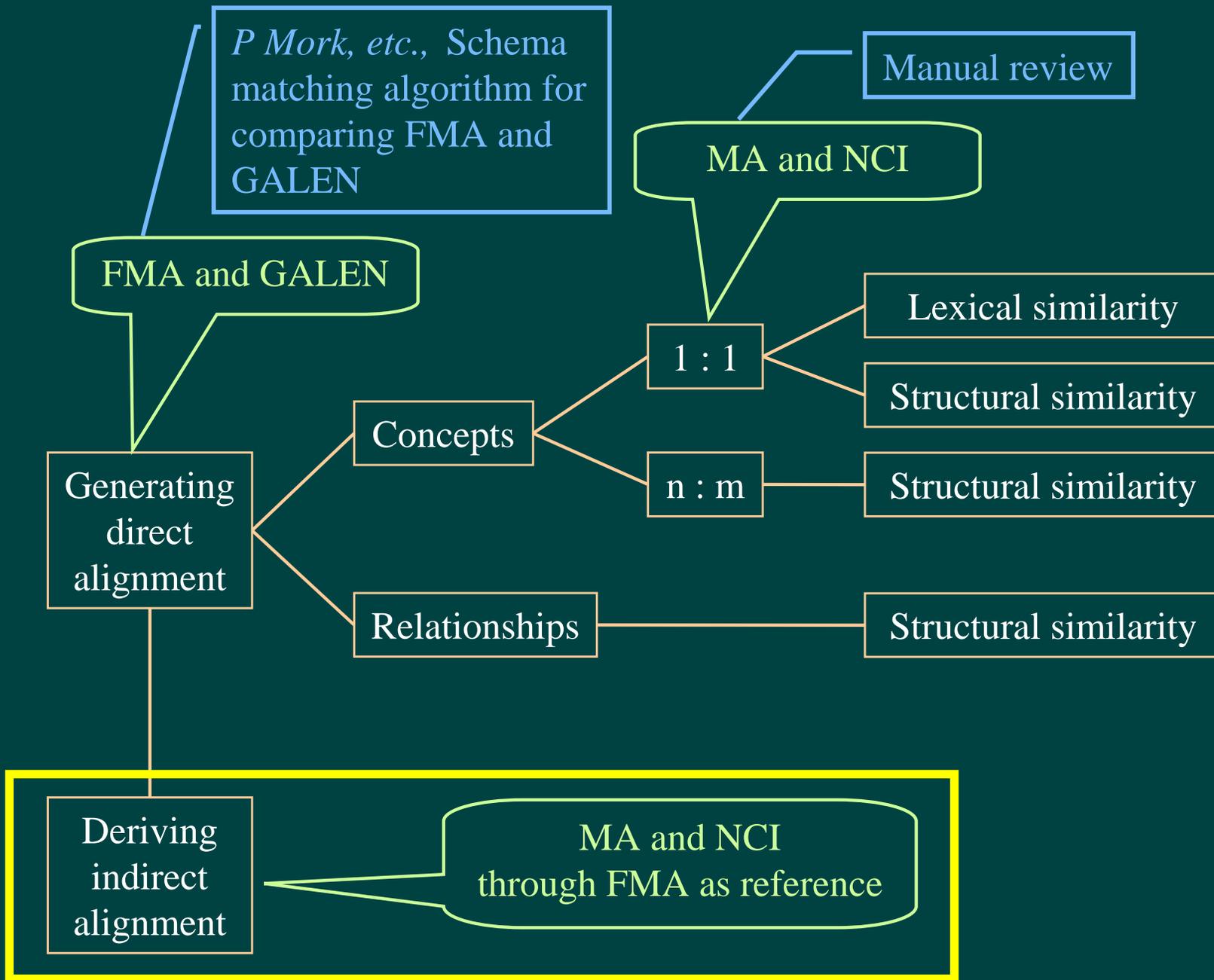
GALEN





Associative relationship mapping patterns

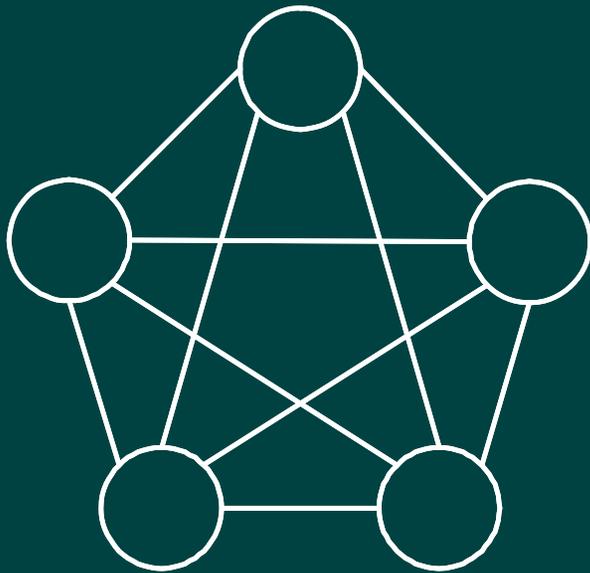
FMA	GALEN	Frequency (N = 4,070)	
<i>branch of</i>	<i>isBranchOf</i>	310	8%
<i>member of</i>	<i>is-a</i>	42	1%
<i>nerve supply</i>	<i>part-of – isServedBy</i>	16	0.4%
<i>part-of – contained in</i>	<i>isNonPartitivelyContainedIn</i>	10	0.25%
<i>contained in</i>	<i>boundsSpace – inverse-isa</i>	2	0.05%



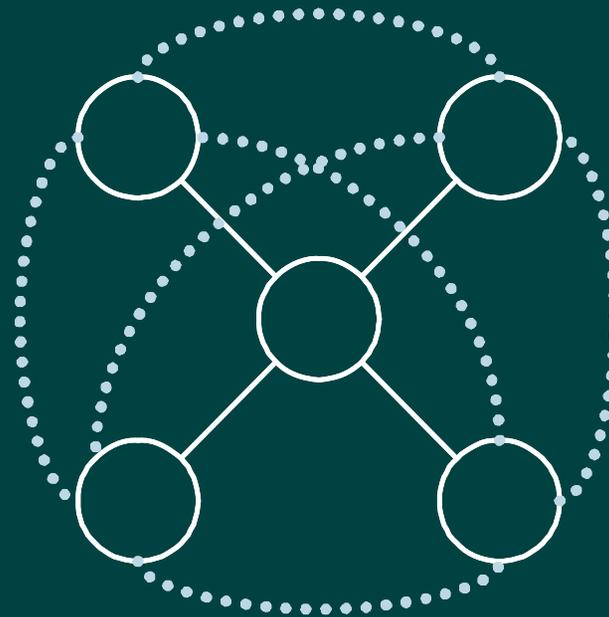
Aligning multiple ontologies of anatomy

Deriving indirect mappings from
direct mappings to a reference

Approaches to aligning multiple ontologies



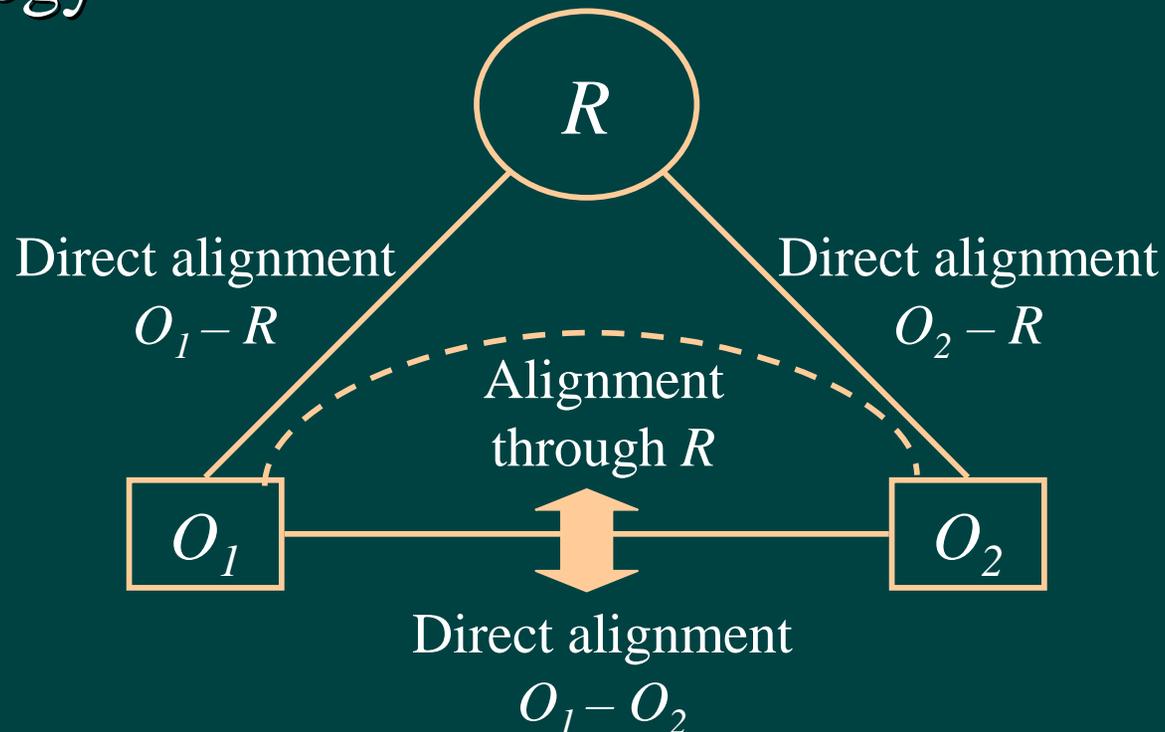
Pairwise
alignment



Alignment through
a reference

Introduction

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



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 identified by 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
 - *Anatomic structure is physical part of*



Foundational Model of Anatomy

- ◆ Conceptualize the physical objects and spaces that constitute the human body
- ◆ Frame-based structure in Protégé
- ◆ 71,202 concepts each identified by 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*

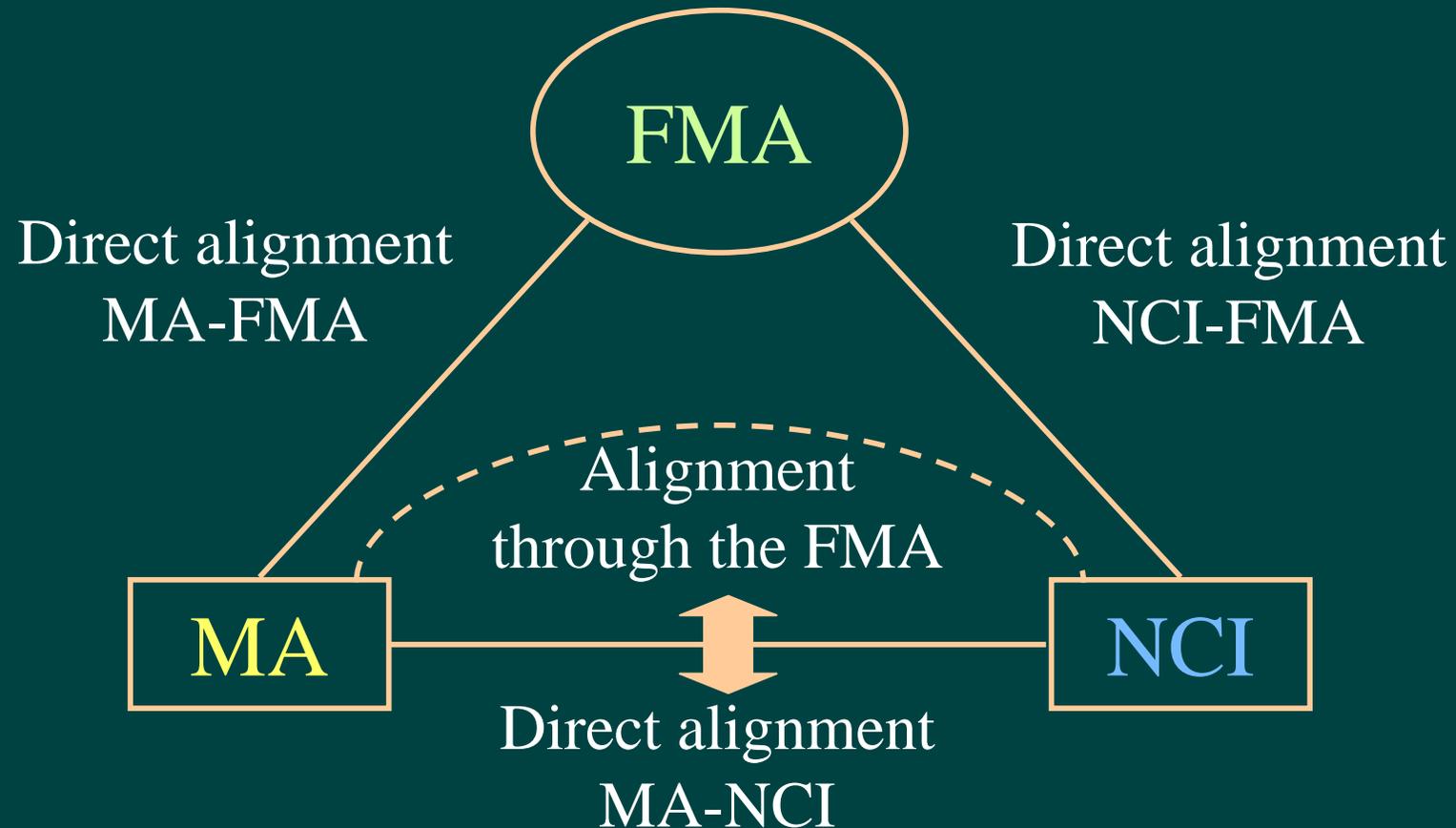


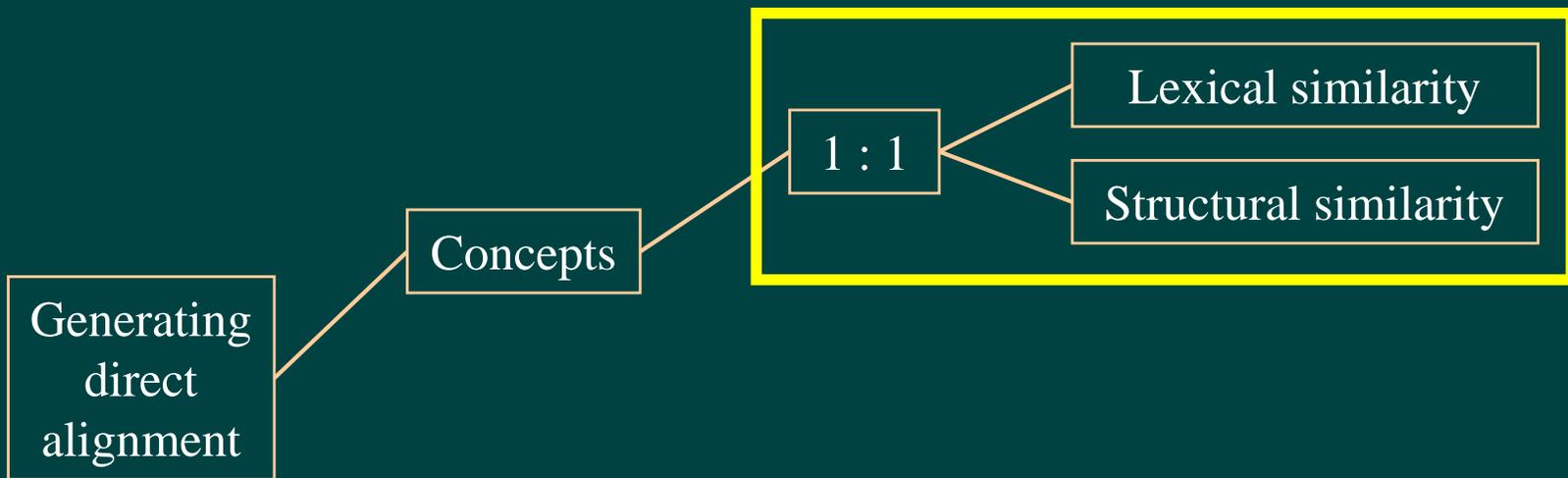
MA, NCI and FMA

	MA	NCI	FMA
Underlying data model	Directed acyclic graph	Available in OWL	Frame-based structure in Protégé
Domain coverage	Mouse anatomy	Human anatomy related to cancers	Human anatomy
Concepts	2,404	4,410	71,202
Synonyms	259	2,371	52,713
Hierarchical relationships	<i>is-a, part-of</i>	<i>is-a, part-of</i>	<i>is-a, part-of (7)</i>
Inverses	-	-	<i>inverse-isa, has-part (7)</i>

Methods

Three phases





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



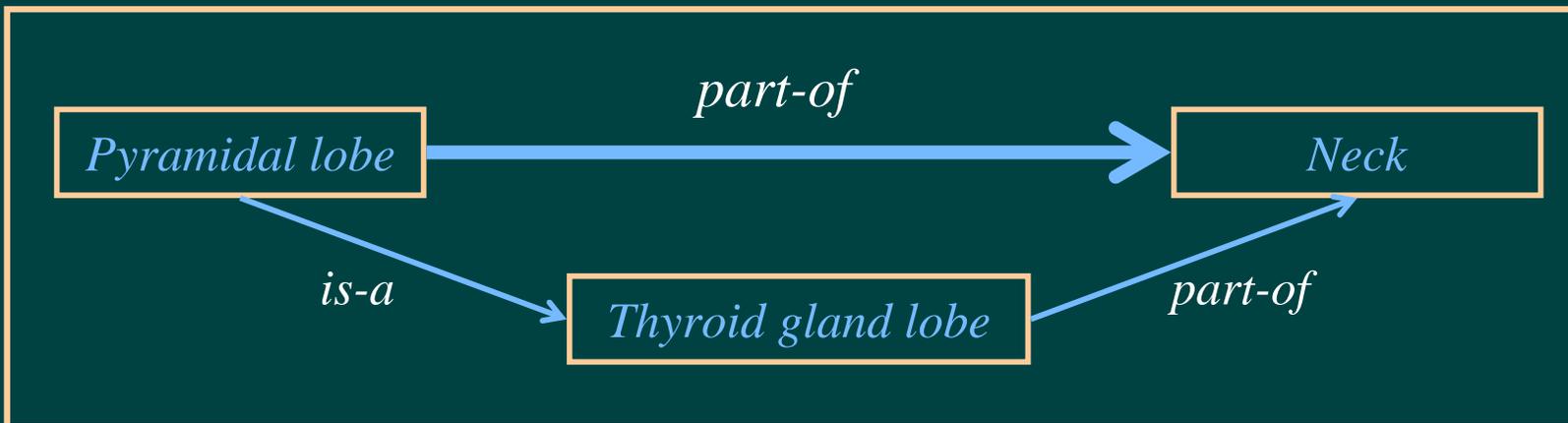
Direct alignment Lexical alignment

- ◆ 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*



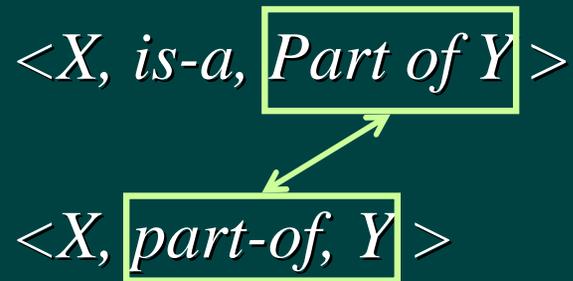
Direct alignment Structural alignment

- ◆ Acquiring inter-concept hierarchical relationships
 - *is-a*, *part-of*, and their inverses *inverse-isa* and *has-part*
- ◆ Complementing missing inverse relations
- ◆ Generating new inter-concept relationships by applying inference rules



Direct alignment Structural alignment

◆ Reification of *part-of* relationships in the FMA

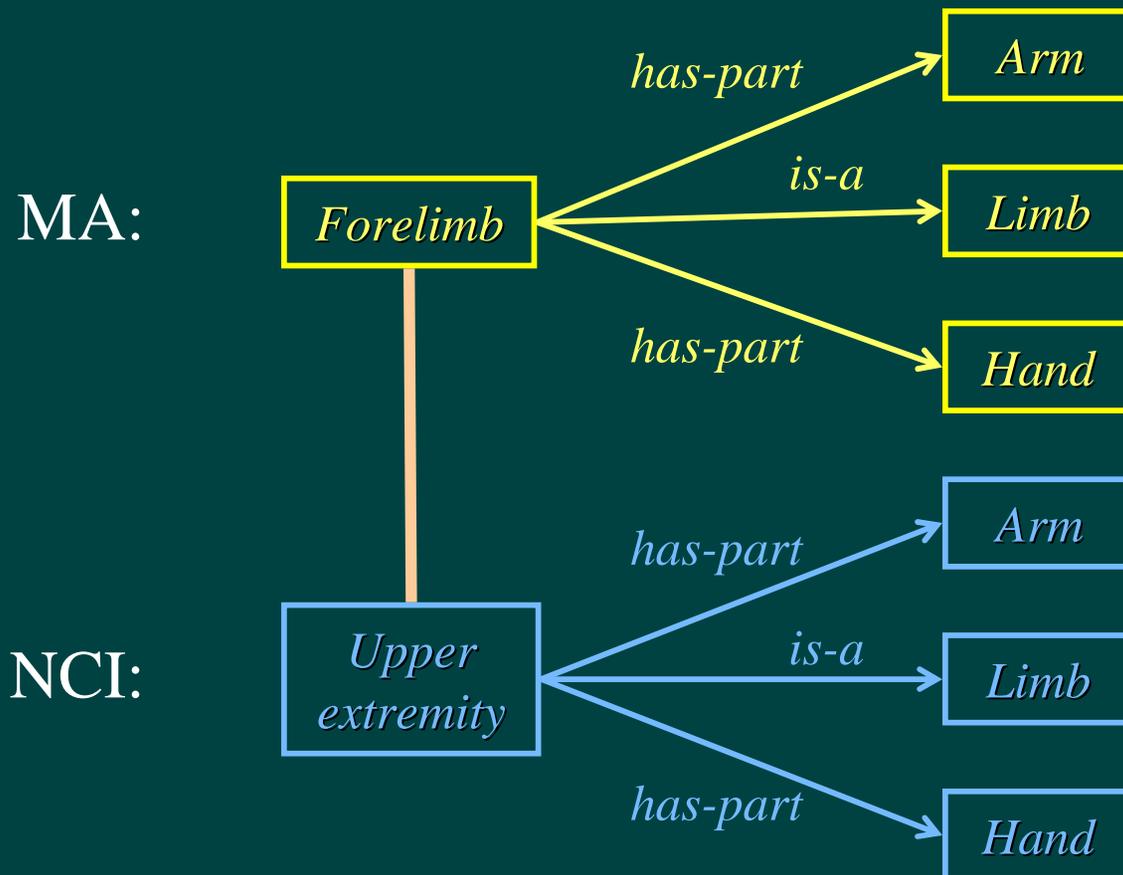


◆ Making explicit reified *part-of* relations in the FMA

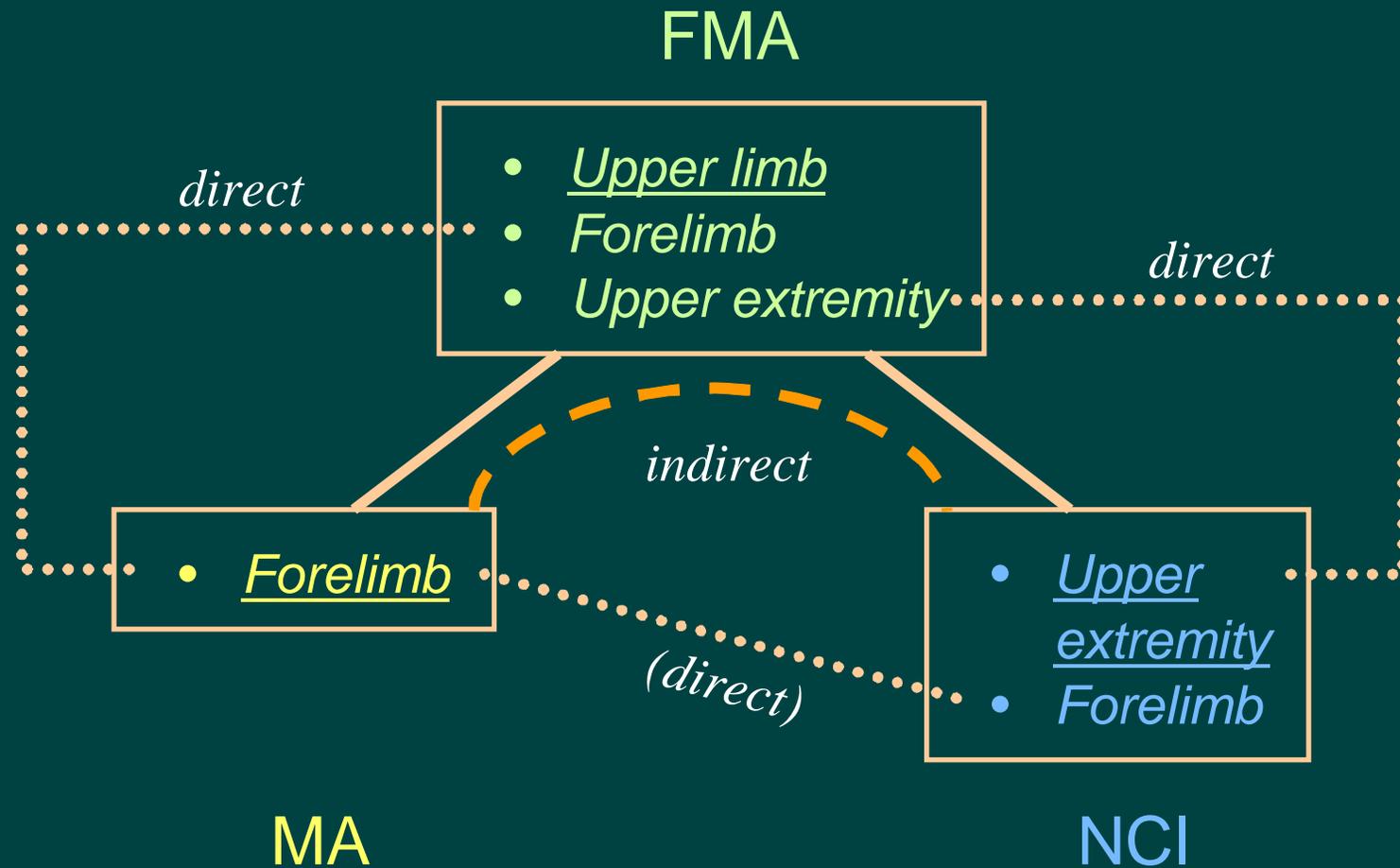
- Reified: $\langle \text{Heel}, \text{is-a}, \text{Subdivision of foot} \rangle$
- Augmented: $\langle \text{Heel}, \text{part-of}, \text{Foot} \rangle$

Direct alignment Structural alignment

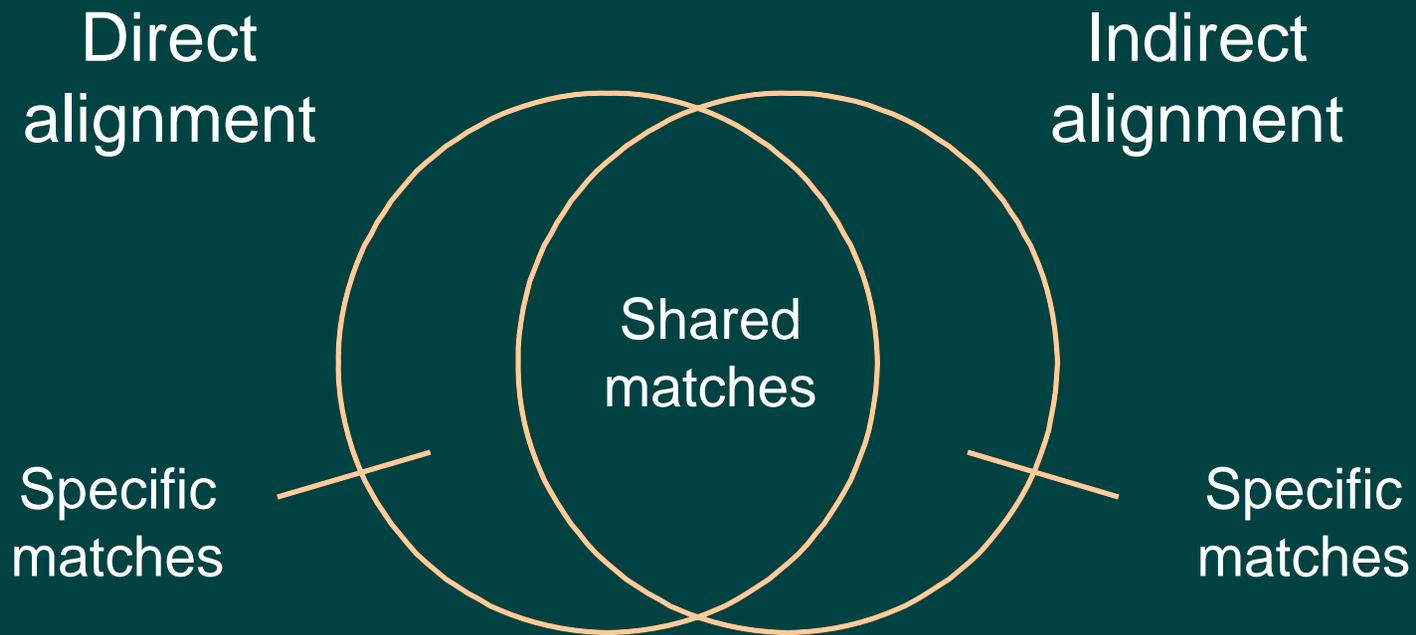
◆ Positive evidence: structural similarity



Phase 2 Indirect alignment



Phase 3 Comparison of two alignments



Results

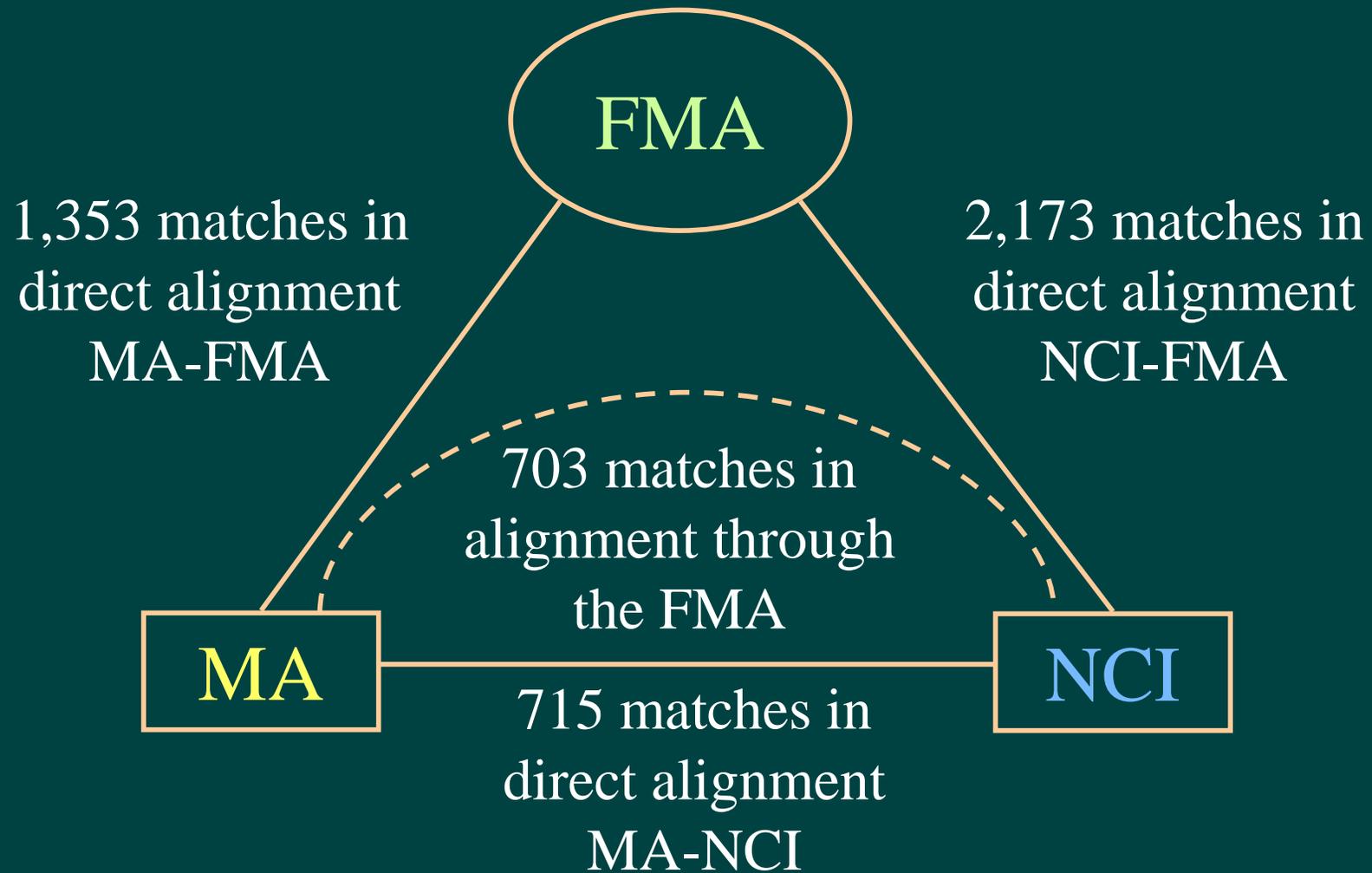
Semantic relations acquired

Types of hierarchical relations	MA	NCI	FMA
Explicitly represented	2,926	7,250	401,045
Complemented	2,926	7,250	4,977
Augmented	-	-	158,282
Inferred	15,044	45,302	5,553,488
Total	20,896	59,802	6,117,792

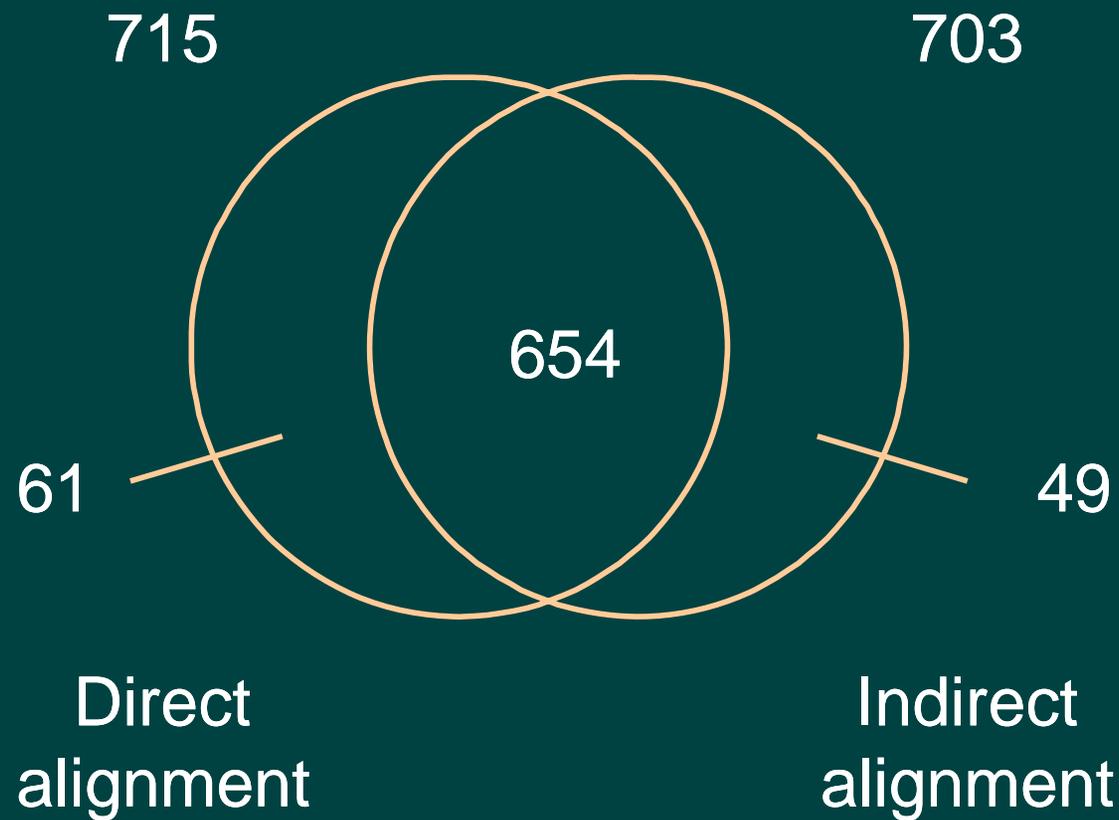
Three direct alignments

	MA - NCI	MA - FMA	NCI - FMA
	715 matches	1,353 matches	2,173 matches
No evidence	62 (8.7%)	66 (4.9%)	205 (9.4%)
Positive evidence	653 (91.3%)	1,283 (94.8%)	1,958 (90.1%)
Negative evidence	0	4 (0.3%)	10 (0.5%)

Indirect alignment



Comparison of two alignments



Comparison of two alignments

◆ 654 shared matches

- 583 (89%) supported by structural evidence in three direct alignments
 - {MA: *Forelimb*, NCI: *Upper extremity*}
- 65 (10%) received no evidence in at least one of the three direct alignments
 - *Chondrocranium* in MA, NCI and FMA
- 6 (1%) received negative evidence in one of the three direct alignments
 - *Pericardial cavity* received negative evidence in MA-FMA
Pericardial cavity received no evidence in MA-NCI
Pericardial cavity received positive evidence in NCI-FMA

Discussion

Benefits of the indirect alignment

- ◆ 49 matches are specific to the indirect alignment between MA and NCI through the FMA
 - 7% of matches in the indirect alignment
 - 45 received positive evidence in both direct alignments MA-FMA and NCI-FMA
 - 2 received no evidence in one of the direct alignments MA-FMA and NCI-FMA
 - 2 received negative evidence in one of the direct alignments MA-FMA and NCI-FMA



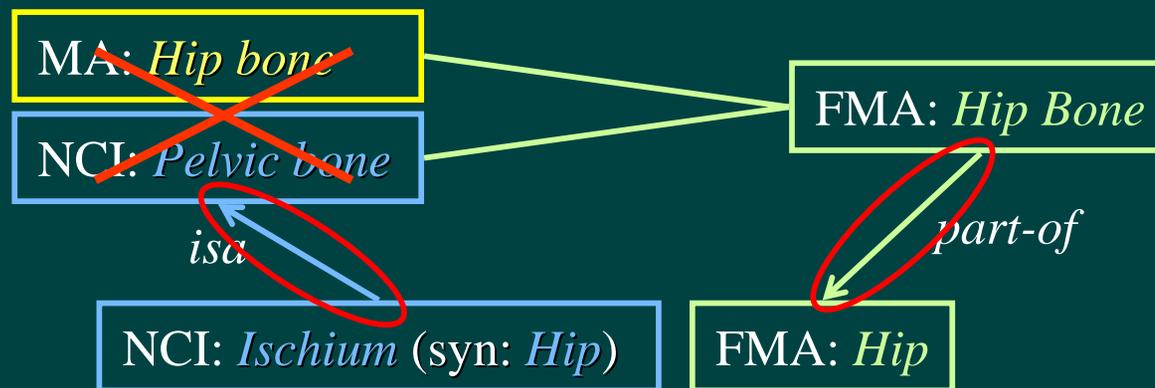
Benefits of the indirect alignment

◆ Why are the 49 matches not identified in the direct alignment?

- Additional synonyms by the FMA



- Additional relations by the FMA



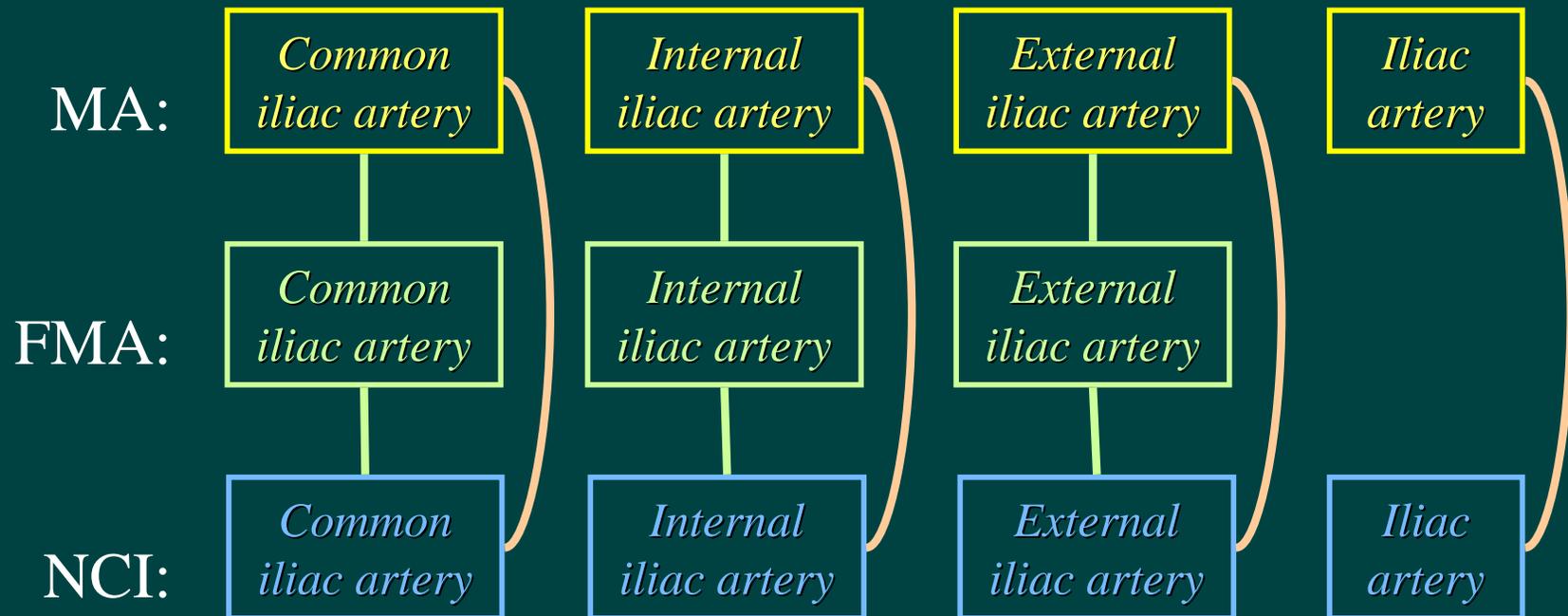
Benefits of the direct alignment

- ◆ 61 matches are specific to the direct alignment
MA-NCI
 - Account for 9% of matches in the direct alignment
 - 53 received positive evidence
 - 8 received no evidence
- ◆ 44 matches identified both directly and indirectly, but received no structural evidence in the indirect alignment
 - 14 received positive evidence in the direct alignment



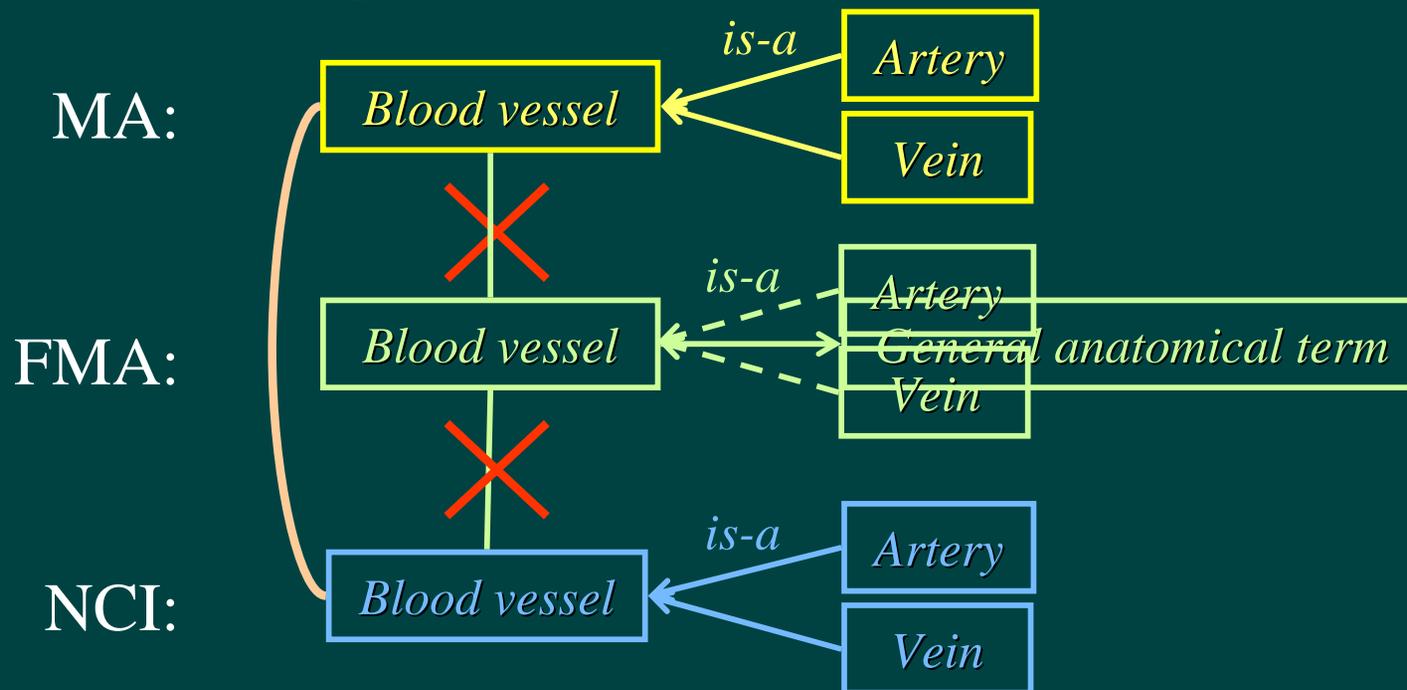
Benefits of the direct alignment

- ◆ Why are the 61 matches not identified in the indirect alignment through the FMA?
 - Different coverage



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



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

Alignment through a reference vs. pairwise alignment

◆ Requirements for a reference ontology

- Broad coverage
- Standard representation principles

◆ The FMA serves as a reference ontology

- Large number of concepts
- Large number of synonyms
- Comprehensive relationships among concepts
- Modeled by a set of declared foundational principles
- Independent of biomedical applications



Current limitations and future work

- ◆ 764 matches identified together by direct and indirect alignments between MA and NCI
 - 32% of MA concepts
 - 17% of NCI anatomical concepts
- ◆ Relies heavily on lexical similarity
- ◆ Only identifies one-to-one concept matches
- ◆ Fully automatic techniques without validation of the alignments
 - No manual validation
 - But cross-validation



Publications

- ◆ Songmao Zhang, Olivier Bodenreider, Knowledge augmentation for aligning ontologies: An evaluation in the biomedical domain. *Proceedings of the Semantic Integration Workshop at the Second International Semantic Web Conference (ISWC 2003)* 2003:109-114
- ◆ Songmao Zhang, Olivier Bodenreider, Aligning representations of anatomy using lexical and structural methods. *Proceedings of AMIA Annual Symposium* 2003:753-757
- ◆ Songmao Zhang, Olivier Bodenreider, Investigating implicit knowledge in ontologies with application to the anatomical domain. In: Altman RB, Dunker AK, Hunter L, Jung TA, Klein TE, editors. *Pacific Symposium on Biocomputing 2004*: World Scientific; 2004:250-261



Publications

- ◆ Songmao Zhang, Peter Mork, Olivier Bodenreider, Lessons learned from aligning two representations of anatomy. In: Hahn U, Schulz S, Cornet R, editors. *Proceedings of the First International Workshop on Formal Biomedical Knowledge Representation (KR-MED 2004)*; 2004:102-108
- ◆ Songmao Zhang, Olivier Bodenreider, Comparing associative relationships among equivalent concepts across ontologies. *Medinfo* 2004:459-463
- ◆ Songmao Zhang, Olivier Bodenreider, Law and Order: Assessing and enforcing compliance with ontological modeling principles. *Computers in Biology and Medicine*; 2005:(accepted)



Publications

- ◆ Songmao Zhang, Peter Mork, Olivier Bodenreider, Philip A. Bernstein, Comparing two approaches for aligning representations of anatomy. *Artificial Intelligence in Medicine*; 2005:(submitted)
- ◆ Songmao Zhang, Olivier Bodenreider, Alignment of multiple ontologies of anatomy: Deriving indirect mappings from direct mappings to a reference. *AMIA*; 2005:(submitted)
- ◆ Olivier Bodenreider, Terry Hayamizu, Martin Ringwald, Sherri De Coronado, Songmao Zhang, Of mice and men: aligning mouse and human anatomies. *AMIA*; 2005:(submitted)
- ◆ Songmao Zhang, Olivier Bodenreider, Frame-based structures vs. description logics in representation and reasoning of the Foundational Model of Anatomy. 2005:(in preparation)
- ◆ Songmao Zhang, Olivier Bodenreider, Structural alignment of anatomical ontologies. 2005:(in preparation)



THANK YOU

