Probabilistic geospatial ontologies

Sumit Sen

sumitsen@uni-muenster.de

11th Nov 2007
Uncertainty Reasoning for the Semantic Web Workshop, BUSAN

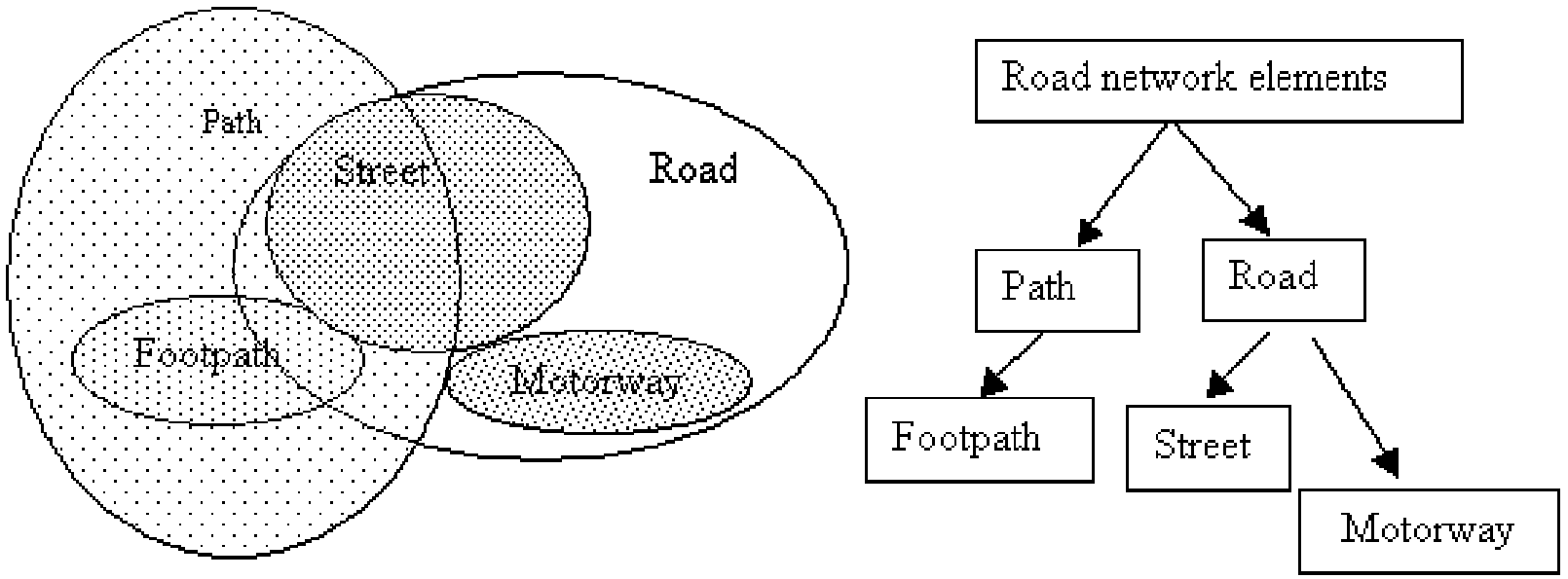
Introduction

- Partial knowledge about geospatial categories is critical for knowledge modeling
- In the geospatial domain
  - Conventional ontologies do not address this issue
- Two components of geospatial ontologies
  - Geospatial Action concepts
  - Geospatial Entity concepts

Overview

- Introduction
  - Road ontology example
  - Need for probabilistic geospatial ontologies
- Case study
  - Highway Code of UK and New York Drivers' Manual ontologies
  - Probabilistic linkages
- BayesOWL ontologies
  - Construction
  - Inferences
- Verification
- Conclusions and future work

Motivation

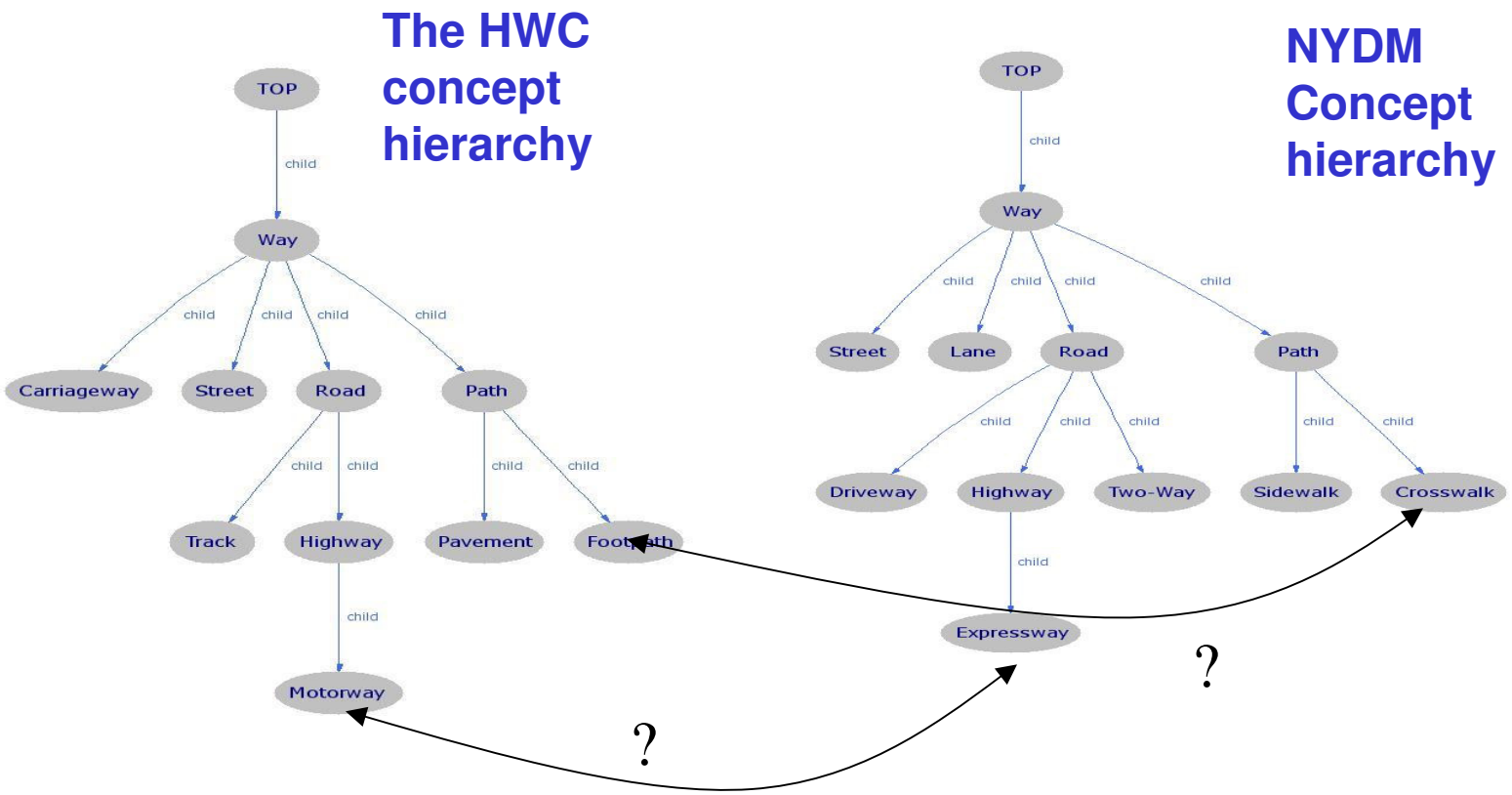


While Highway and Street are subclasses of Road, Footpath is a subclass of Path. Evidently this representation shows that Highway and Footpath are small subclasses of Road and Path respectively. Street has a major overlap with Path allthough it is not a subclass. Representation of the five classes as a subsumption relation in a conventional ontology does not represent or allow inferences based on the additional information about overlaps

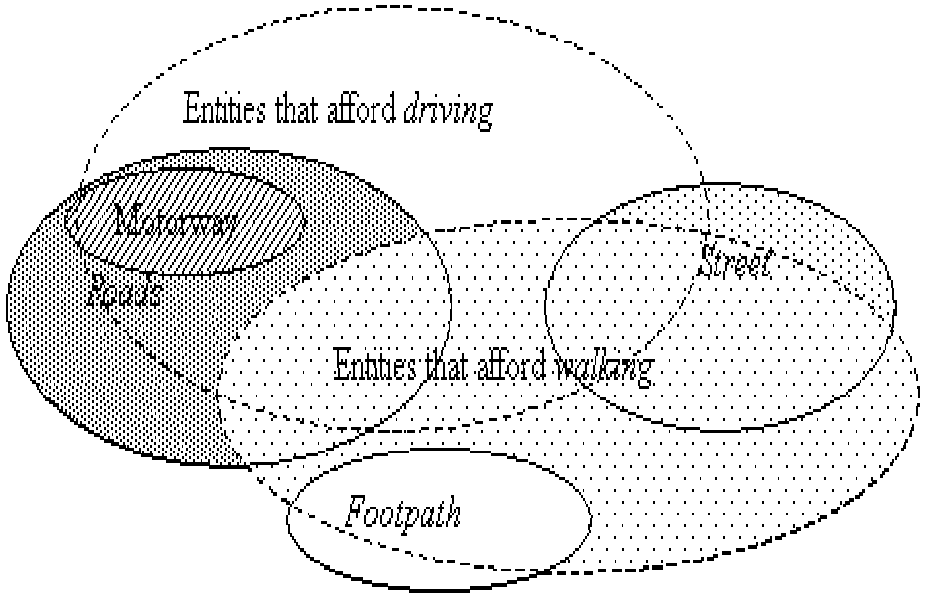
Motivation …

Concepts about actions and entities and why we need linkages

- Most similar entity for a Motorway, Footpath
- Degree of similarity



Motivation …



Representaion of overlaps between some entity concepts and action based concepts for road networks in the UK.

While ellipses with solid borders represent geospatial entities, the ellipses with dashed borders represent abstract concepts based on the entities that afford certain geospatial action.

What is a Motorway?

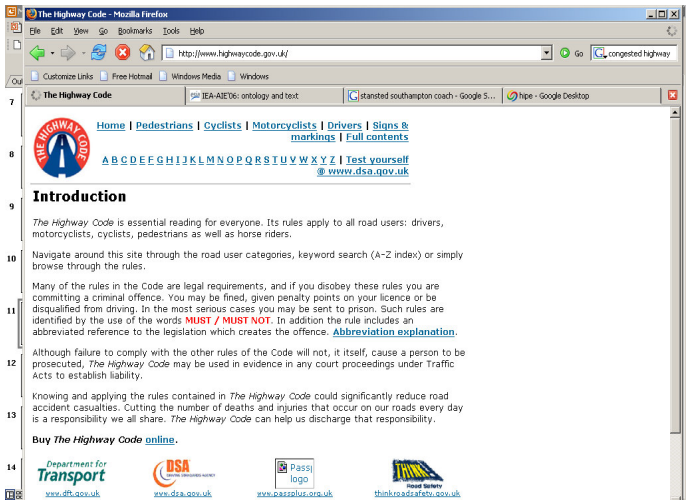
What is a Footpath?

Case study -Methodology

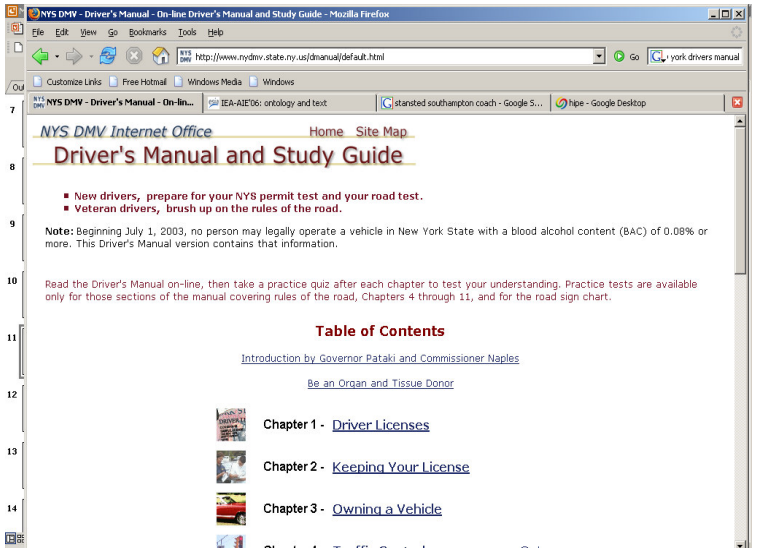
- Ontology Extraction
  - Lexicographic and frequency analysis
  - Hierarchies based on is-a relation of both entities and actions
- Reasoning and Inferences (most similar and most dissimilar concepts)
  - Within a single ontology
  - Across ontologies

Case study

The Highway Code:



The New York Driver's Manual



Tagged text

#	Token	Lemma	POS	Chunk	Relation	Sense	Sense Definition
179	You	you	PRP	NP-B	NPSBJ-B	no-sense	
180	must	must	MD	VP-B	VP-B	no-sense	
181	come	come	VB	VP-I	VP-I	come%2:38:04::	reach a "destination " arrive by movement or by making "progress " "She arrived home at 7 o'clock " "He got into college " She didn't get to Chicago until after midnight
182	to	to	TO	PP-B	PNP-B	no-sense	
183	a	a	DT	NP-B	PNP-I	no-sense	
184	stop	stop	NN	NP-I	PNP-I	stop%1:11:00::	the event of something "ending " it came to a stop at the bottom of the hill
185	before	before	IN	PP-B	PNP-B	no-sense	
186	the	the	DT	NP-B	PNP-I	no-sense	
187	stop	stop	NN	NP-I	PNP-I	stop%1:11:00::	the event of something "ending " it came to a stop at the bottom of the hill
188	line	line	NN	NP-I	PNP-I	line%1:06:00::	something (as a cord or rope) that is long and thin and "flexible " a washing line
189	,	,	,	,	,	no-sense	

GAMBL output: NYDM 179-189

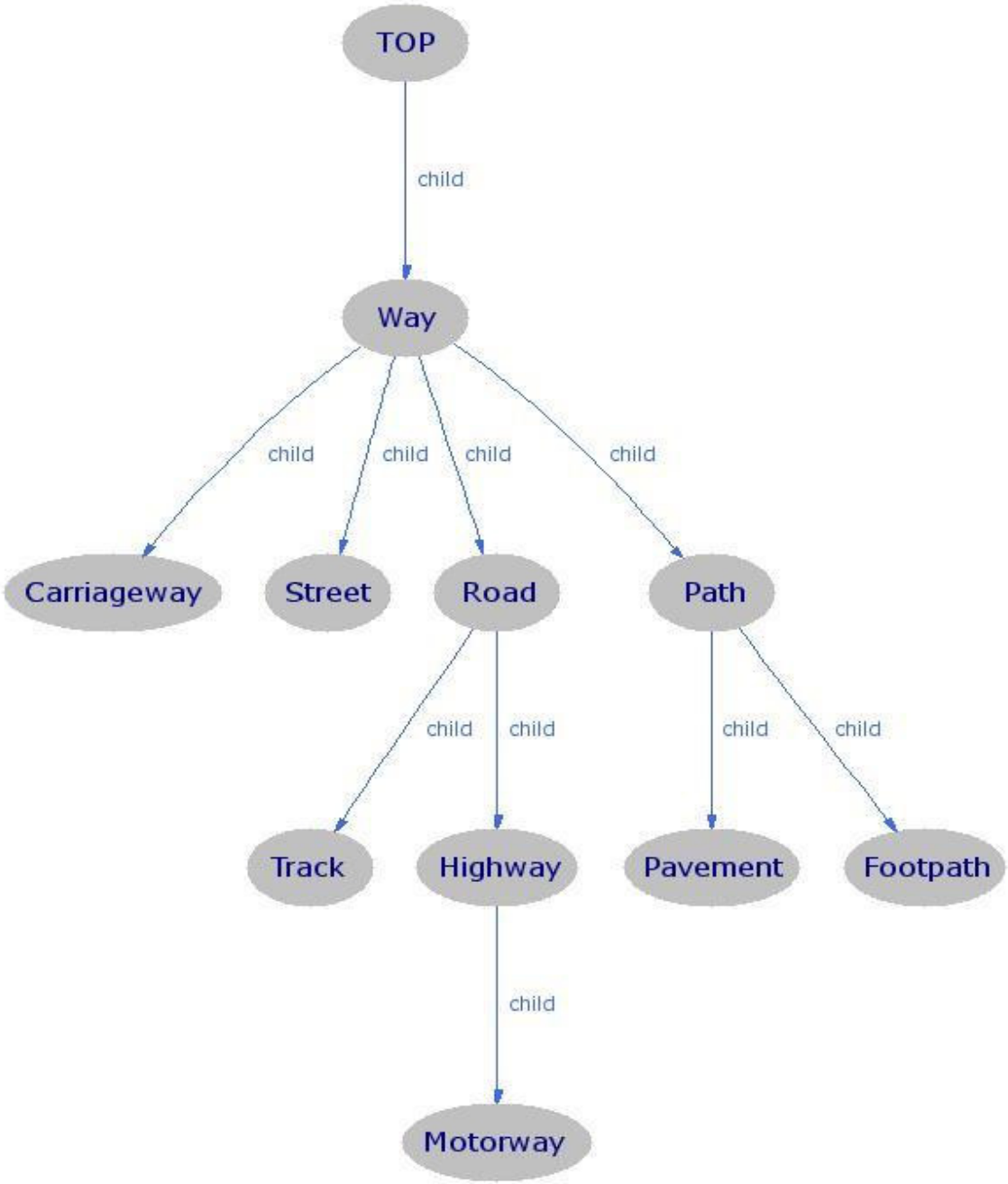
Extracting nouns - NYDM

Term	*Sense*	*Sense definition*
Driveway	driveway%1:06:00::	a road leading up to a private "house " they parked in the driveway
Road	road%1:06:00::	an open way (generally public) for travel or transportation
Lane	lane%1:06:00::-(default)	a narrow way or road
Way	way%1:04:01::	how a result is obtained or an end is "achieved " "a means of control " "an example is the best agency of instruction " the true way to success
Crosswalk	crosswalk%1:06:00::	a path (often marked) where a street or railroad can be crossed
Two-way(road)	two-way%5:00:00:bidirectional:00-(default)	operating or permitting operation in either of two opposite "directions " "a two-way valve " "two-way traffic " two-way streets
Street	street%1:06:00::	a thoroughfare (usually including sidewalks) that is lined with buildings; they walked the streets of the small town; he lives on Nassau Street
U-turn	u-turn%1:04:00::	complete reversal of direction of travel
Path	path%1:04:00::	a course of conduct; the path of virtue; we went our separate ways; our paths in life led us apart; genius usually follows a revolutionary path
Route	route%1:15:00::	an established line of travel or access
Incline	incline%1:06:00::-(default)	an inclined surface or roadway that moves traffic from one level to another or axle (as in vehicles or other machines)
Expressway	expressway%1:06:00::	a broad highway designed for high-speed traffic
Sidewalk	sidewalk%1:06:00::	walk consisting of a paved area for "pedestrians " usually beside a street or roadway

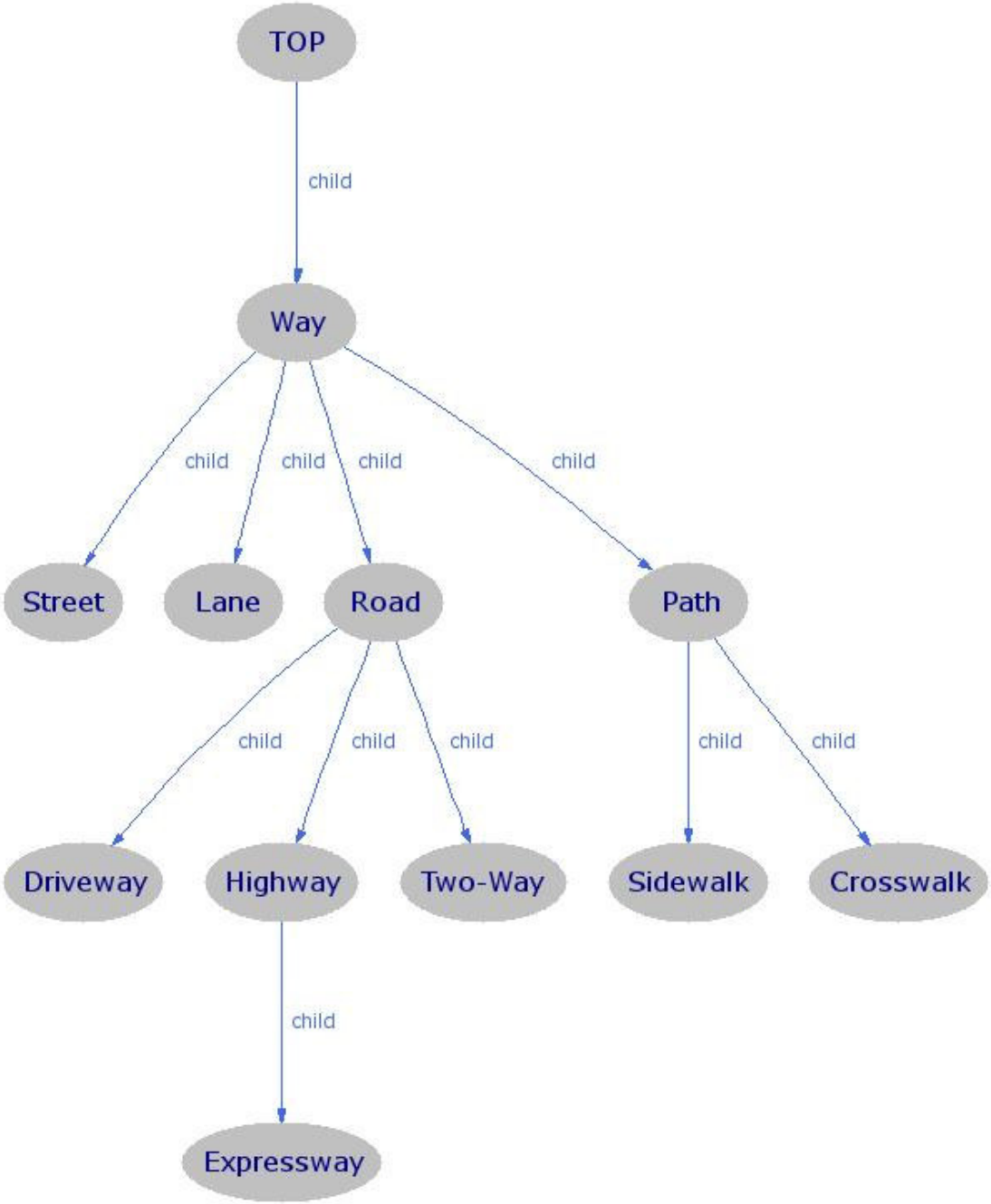
Nouns from the HWC

Term	*Sense*	*Sense definition*
Motorway	motorway%1:06:00::	a broad highway designed for high-speed traffic
Road	road%1:06:00::	an open way (generally public) for travel or transportation
Carriageway	carriageway%1:06:0(	(British) one of the two sides of a motorway where traffic travels in one direction only usually in two or three lanes
Footpath	footpath%1:06:00::	a trodden path
Street	street%1:06:00::	a thoroughfare (usually including sidewalks) that is lined with buildings; they walked the streets of the small town; he lives on Nassau Street
Pavement	pavement%1:06:00::	the paved surface of a thoroughfare
Footbridge	footbridge%1:06:00::	a bridge designed for pedestrians
Kerb	kerb%1:06:00::	an edge between a sidewalk and a roadway consisting of a line of curbstones (usually forming part of a gutter)
Path	path%1:04:00::	a course of conduct; the path of virtue; we went our separate ways; our paths in life led us apart; genius usually follows a revolutionary path
Lane	lane%1:06:00::-(defa	a narrow way or road

The HWC concept hierarchy

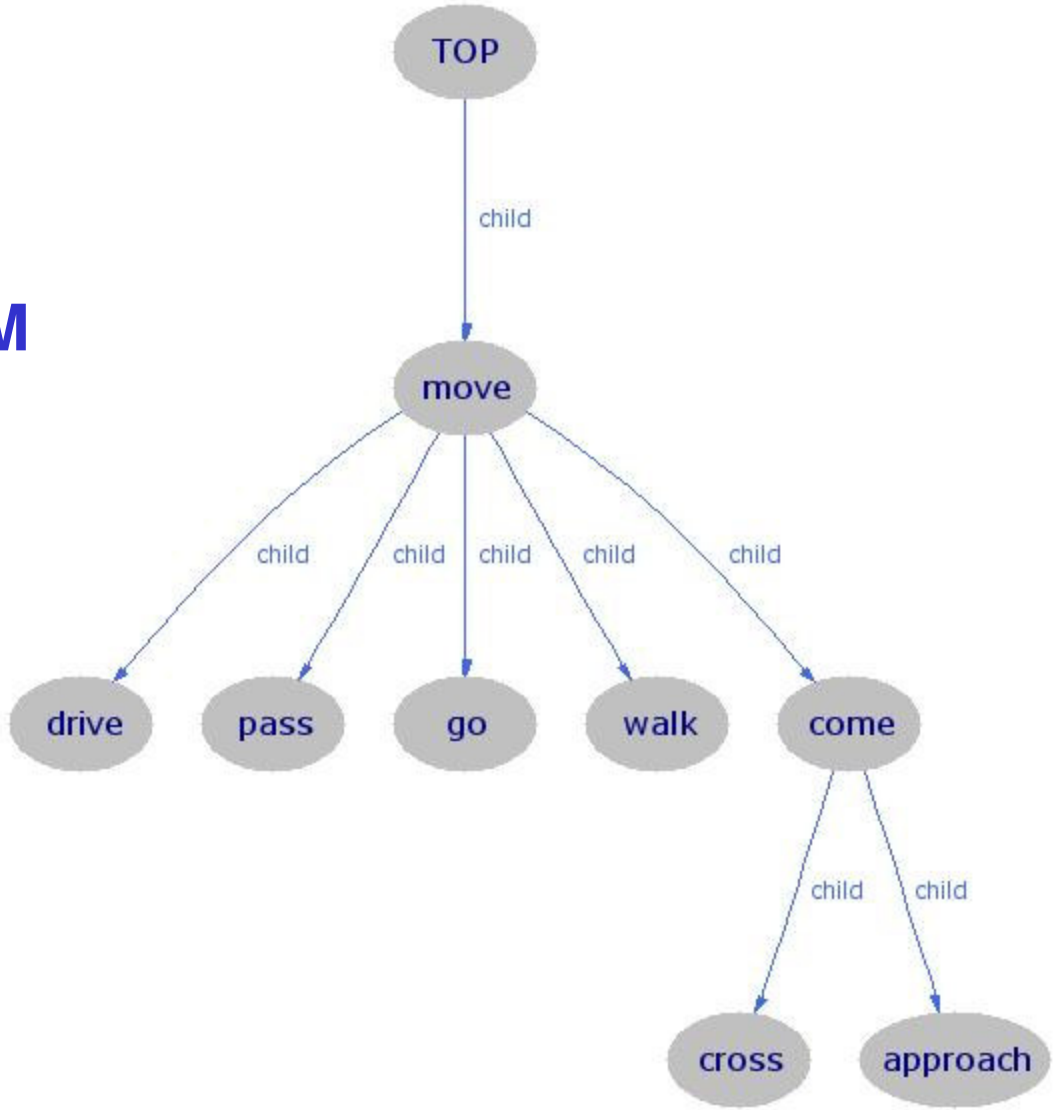


NYDM Concept hierarchy



Action/function hierarchy

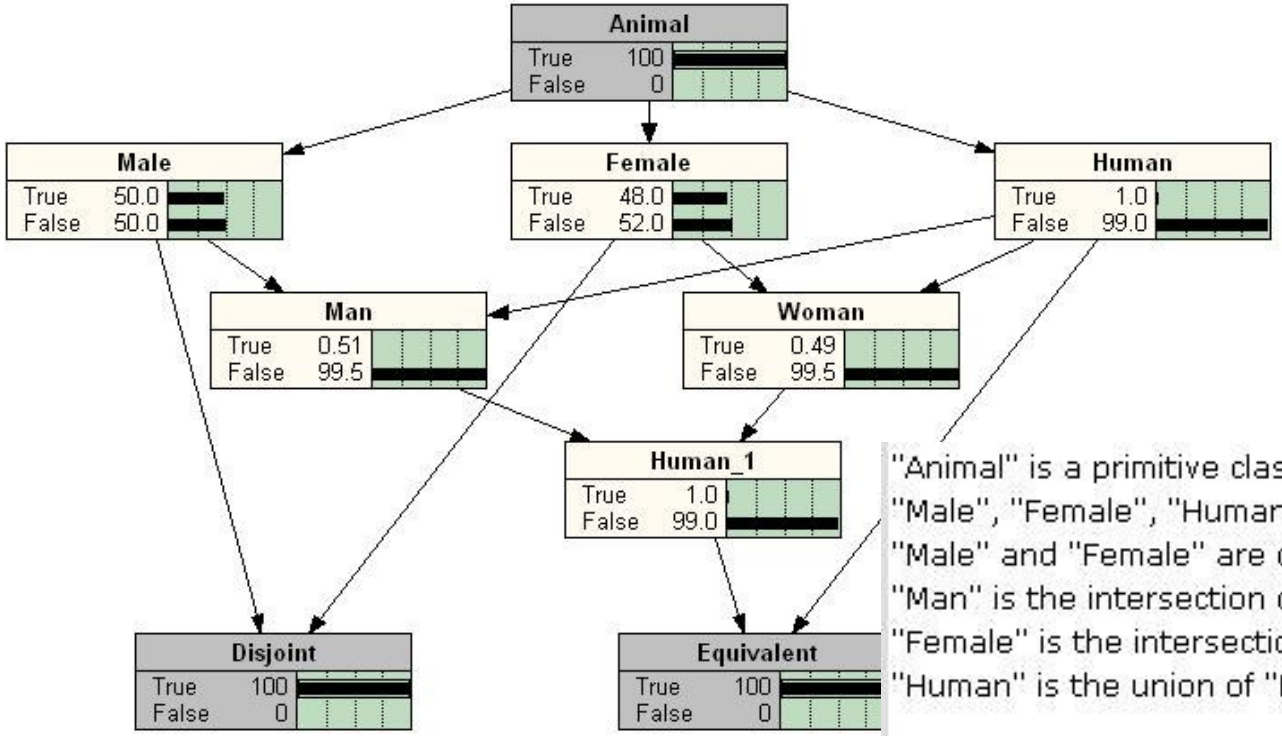
Both HWC and NYDM



Linkages: Entities-functions

NYDM	Street	Road	Footpath	Motorway	Lane	Way	Path	Crosswalk	Expresswa
move	0.015	0.049		0.012	0.107	0.035	-	-	-
walk	-	0.026	0.056	0.000	-	-	-	-	-
drive	0.057	0.062	0.000	0.069	0.000	-	-	-	-
enter	-	0.025	-	-	0.000	0.020	-	-	-
stop	0.010	0.075	-	0.000	0.000	0.051	-	-	-
be	0.014	0.215	0.006	0.028	0.061	0.033	0.014	-	-
cross	0.029	0.135	-	0.000	0.024	0.067	0.020	-	-
turn	0.038	0.059	-		0.042	0.041	-	-	-
wait	-	0.040	-	0.000	0.009	0.031	-	-	-
approach	0.022	0.052	-	0.016	0.065	0.045	0.023	-	-
go	-	0.021	-	-	0.063	-	-	-	-
pass	-	0.038	-	-	0.032	0.012	0.017	-	-
HWC									
move	0.026	0.032	-	-	0.107	-	0.032	-	-
walk	-	0.010	-	-		-		-	-
drive	0.020	0.061	-	-	0.056	-	-	-	0.047
enter	0.025	0.048	-	-	0.077	0.041	-	0.053	0.064
stop	0.019	0.048	-	-	0.038	0.026	-	0.059	0.026
be	0.011	0.068	-	-	0.089	0.026	0.004	0.009	0.024
cross	0.061	0.033	-	-	0.017	0.071	-	0.030	
turn	0.037	0.080	-	-	0.094	0.051	0.029	0.018	0.008
wait	0.040	-	-	-	0.009	0.059	-	-	0.029
approach	0.015	0.060	-	-	0.034	-	-	-	0.026
go	0.020	0.029	-	-	0.030	0.051	-	-	0.017
pass	0.044	0.039	-	-	0.130	0.025	-	0.014	0.013

BayesOWL- A mild intro



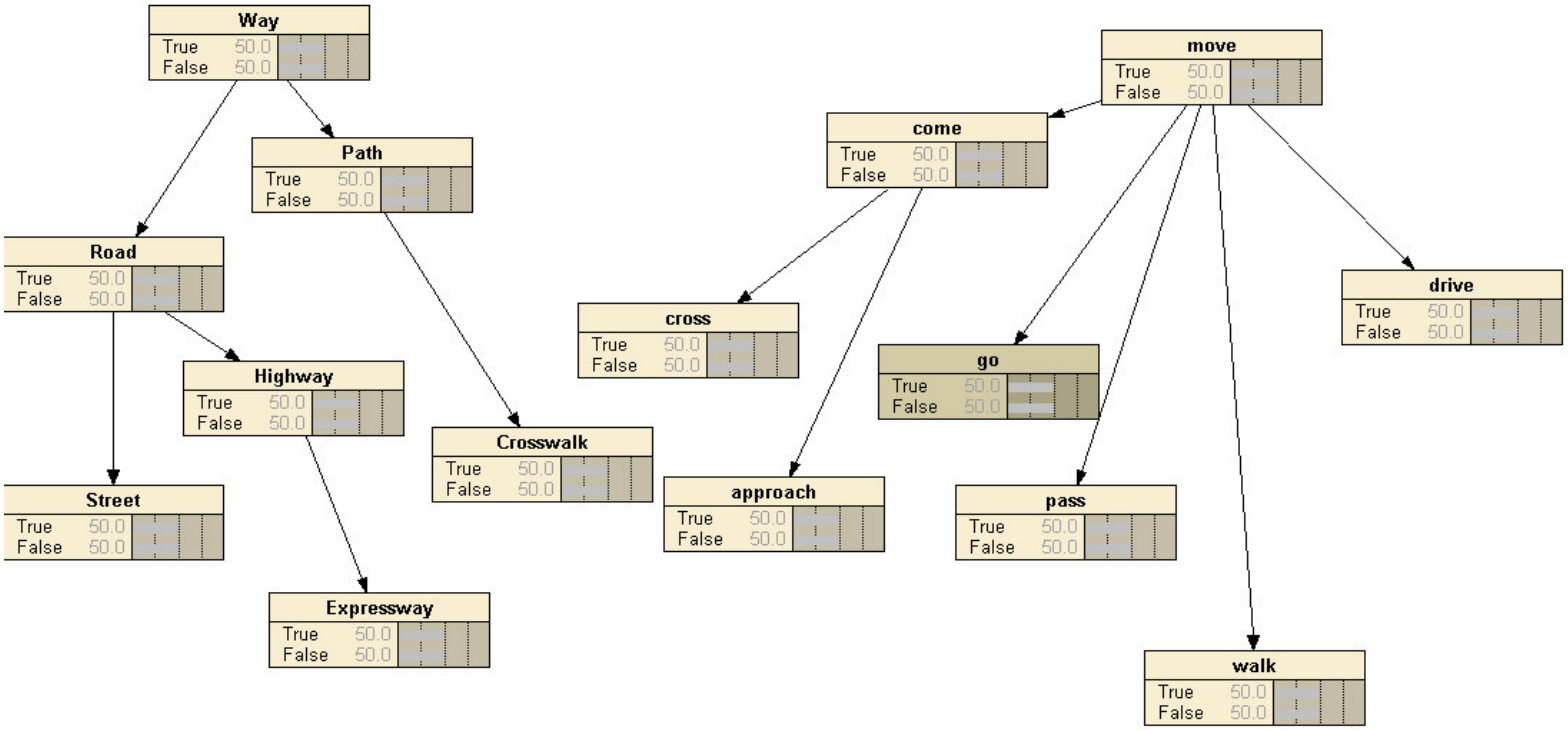
"Animal" is a primitive class
"Male", "Female", "Human" are subclasses of "Animal"
"Male" and "Female" are disjoint with each other
"Man" is the intersection of "Male" and "Human"
"Female" is the intersection of "Female" and "Human"
"Human" is the union of "Man" and "Woman"

Probability information:

P(Animal) = 0.5
P(Male|Animal) = 0.5
P(Female|Animal) = 0.48
P(Human|Animal) = 0.01

Source: Ding *et al* 2005

BayesOWL for Geospatial ontologies



Inferences within an ontology

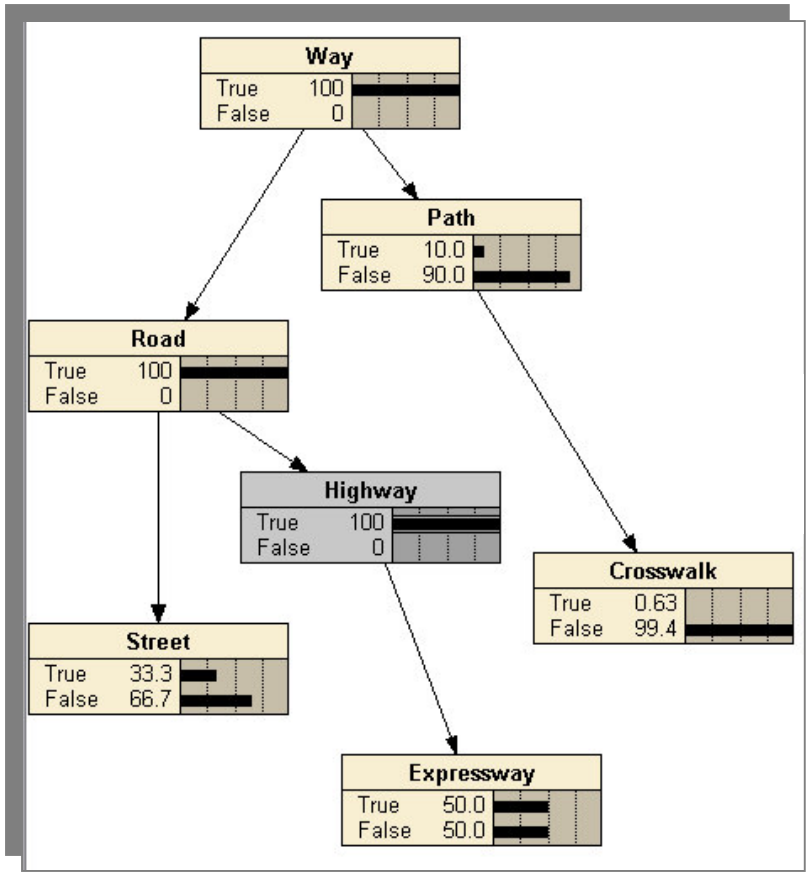
Entity Concept	Occurs in	Most similar function concept		Most dissimilar function concept	
Crosswalk	NYDM	cross		move,go	
Expressway	NYDM	drive		cross	
Footpath	HWC	cross		drive	
Highway	NYDM/HWC	drive	drive	walk	go,move
Motorway	HWC	drive		cross,walk	
Path	NYDM/HWC	move,go	cross	cross	move,go
Road	NYDM/HWC	drive	drive	cross,walk	cross,walk
Street	NYDM/HWC	cross,walk	cross,walk	go	go
Way	NYDM/HWC	move,go	move,go	cross	cross,walk

Probabilistic inferences

- Simple adaptations of BayesOWL

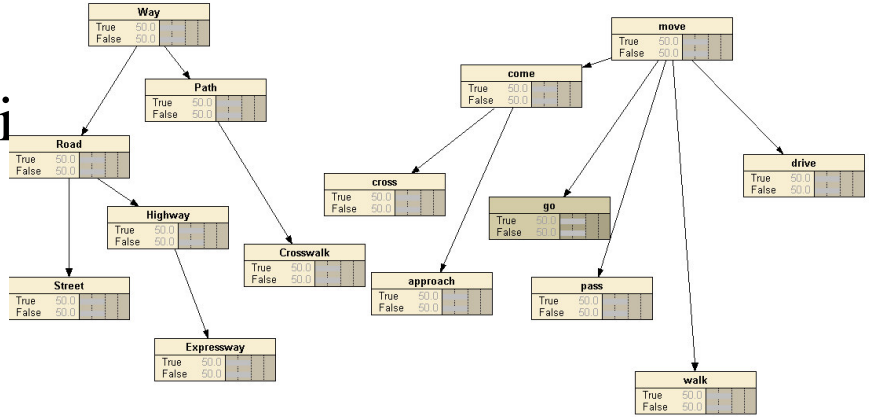
(without $\neg$, $\wedge$, $\vee$, $\equiv$)

- Some inferences
  - Equivalent concepts and extent of overlap
  - Most Similar and dissimilar concepts
  - Inconsistent concepts



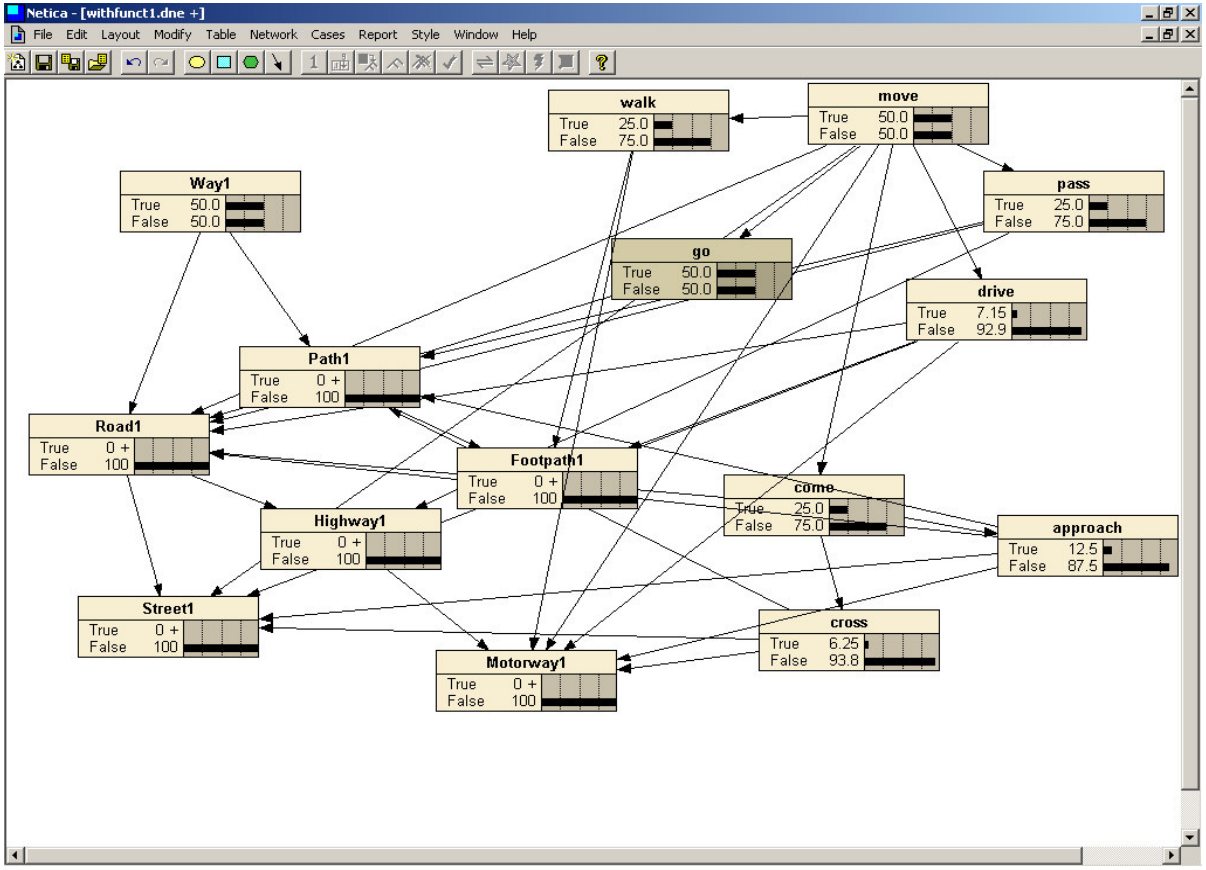
Linking the BNs

- Virtual evidence via functi
- Labels are not important
  - Functions are

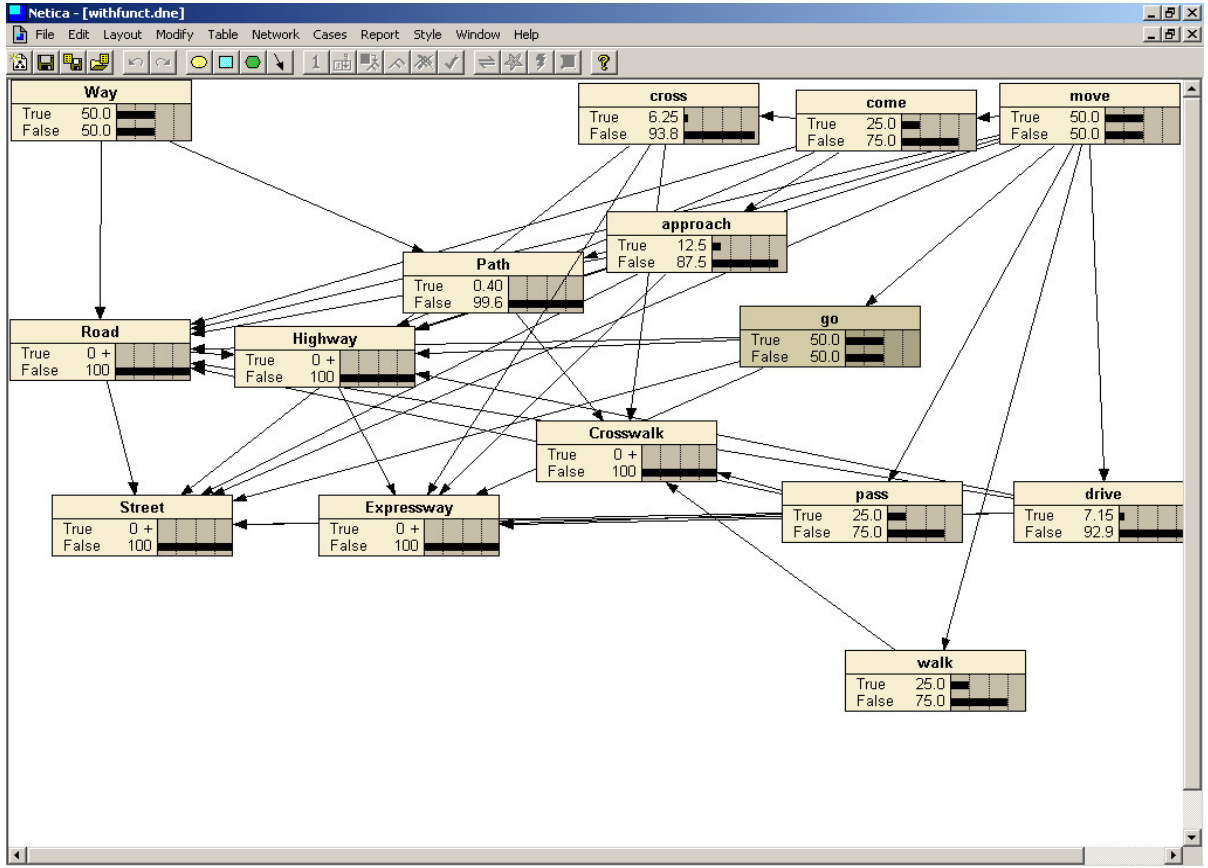


HWC	NYDM	pass	go	drive	walk	cross	approach	come	move
Footpath1	**Street**	0.044	0.020	0.020	**0.001**	0.061	0.015	**0.001**	0.026
Footpath1	**Road**	0.039	0.029	0.061	**0.001**	**0.001**	0.060	0.026	0.032
Footpath1	**Way**	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**
Footpath1	**Path**	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**	0.032
Footpath1	**Crosswalk**	0.014	**0.001**	**0.001**	0.010	0.030	**0.001**	**0.001**	**0.001**
Footpath1	**Expressway**	0.013	0.017	0.047	**0.001**	**0.000**	0.026	**0.001**	**0.001**

Linking the hierarchies-HWC



…NYDM


Netica - [withfunct.dne]
File Edit Layout Modify Table Network Cases Report Style Window Help
Way
True 50.0
False 50.0
cross
True 6.25
False 93.8
come
True 25.0
False 75.0
move
True 50.0
False 50.0
approach
True 12.5
False 87.5
Path
True 0.40
False 99.6
Road
True 0 +
False 100
Highway
True 0 +
False 100
go
True 50.0
False 50.0
Crosswalk
True 0 +
False 100
Street
True 0 +
False 100
Expressway
True 0 +
False 100
pass
True 25.0
False 75.0
drive
True 7.15
False 92.9
walk
True 25.0
False 75.0

Inferences across ontologies

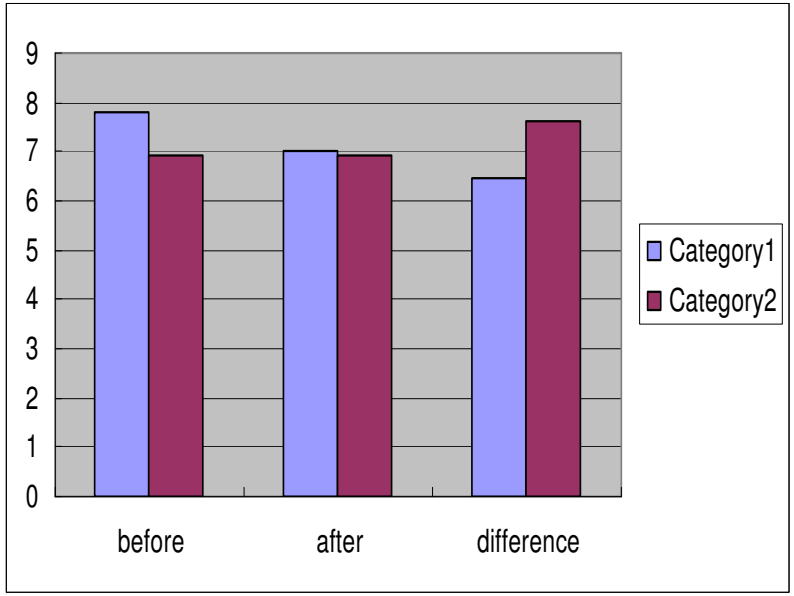
HWC Concept	Most similar entity	Most dissimilar entity
Footpath	Path	Expressway
Highway	Way	Street
Motorway	Road	Crosswalk
Path	Path	Expressway
Road	Road	Expressway
Street	Path	Street
Way	Way	Expressway

NYDM Concept	Most similar entity	Most dissimilar entity
Way	Way	Motorway
Street	Way	Street
Road	Road	Street
Path	Path	Motorway
Highway	Path	Street
Expressway	Road	Street
Crosswalk	Path	Motorway

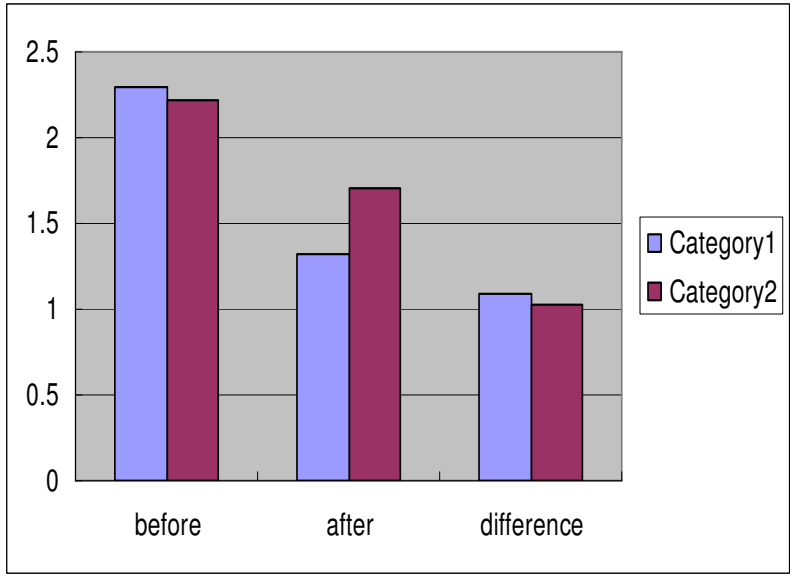
Human subjects testing

- 20 Subjects
  - Both Genders
  - Only familiar with HWC and familiar with both HWC and NYDM
  - Ages 25 to 60
- Game of cards
  - Matching concepts
    - Without entity names
    - Entity names with descriptions of actions permitted on the entity
- Even distribution of results

Changes in mappings are consistent for both categories



Average of number of mappings



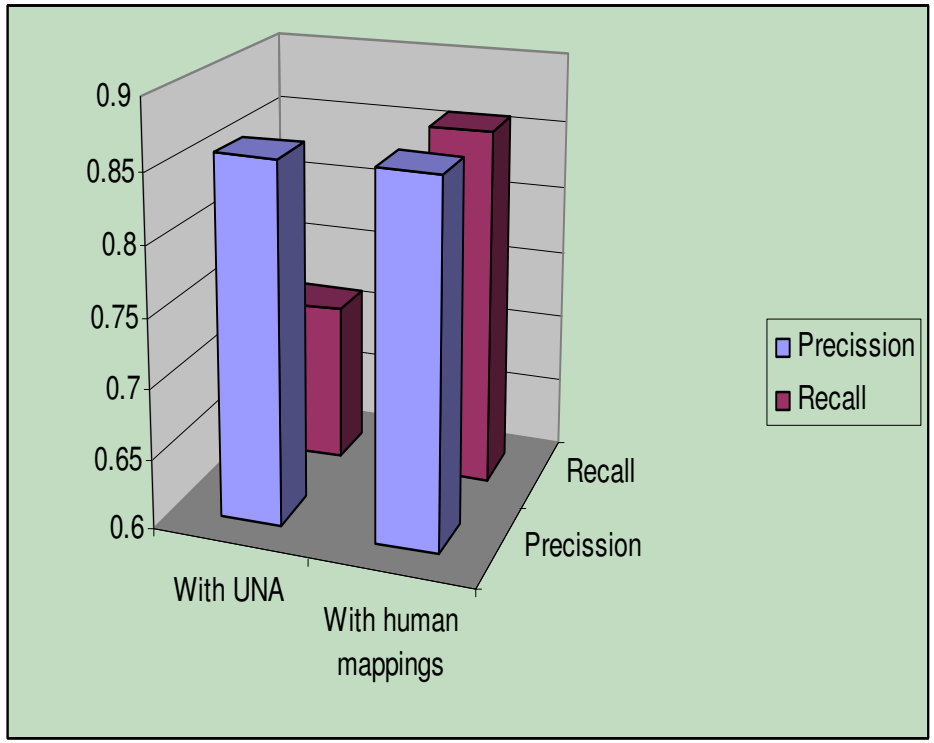
Standard Deviation of number of mapping by the subjects

Analysis

•Close resemblance between machine based matches and those from human subjects testing

•Entity names do not mean exactly the same; quantification of the differences is possible if actions are assumed invariant

•Human perception of the meaning of entity names change when the actions afforded by the entity is suggested to be different



Main Conclusions

- Ontologies of geospatial entities need to be extended with probabilistic frameworks
- It is possible to use both hierarchies of geospatial entities as well as geospatial actions and link them with probabilistic knowledge about affordances of geospatial entities.
- The use of probabilistic geospatial ontologies for mappings between most similar entities mimics, to a large extent, the human mechanism of semantic translations of entity names.
- Our results provide support to the hypothesis that knowledge about geospatial actions and affordances to such actions are a critical part of geospatial knowledge.

Directions for future work

- **Inclusion of Disjoint, Equivalent, Intersection and Union relations:** Using such relations in future will require use of some iterative algorithm such as Decomposed IPFP in order to enforce truth conditions of the LNodes in BayesOWL [Ding *et al,* 2005].
- **Testing on industrial scale:** this experiment, although at a prototype scale aims, in the end, to solve semantic problems, which occur at industrial scale.
- **Machine based learning:** The human mappings, especially that of the experts, are considered as the ideal mappings. Human interactions and judgments for most similar concepts can be used to improve heuristics involved in specification of entity-action linkages.

Thank you

Sumitsen@uni-muenster.de