

# WordNet and Urban Civil Engineering

## Preliminary study

G. Falquet

C21 - WG3 meeting  
May 2006

# Contents

## WordNet

- objectives
- structure and content

## UCE Concepts in WordNet

- WN and the Urbamet thesaurus
- Observations

## WordNet and UCE applications

# WordNet: a Lexical Ontology

## Purposes:

- to produce a combination of dictionary and thesaurus that is intuitively usable
- to support automatic text analysis and artificial intelligence applications.

Released under a BSD style license and can be downloaded and used freely.

The database can also be browsed online

<http://wordnet.princeton.edu/perl/webwn2.1>

## Concepts = synsets

Information in WordNet is organized around logical groupings called **synsets** (synonym sets).

Each synset consists of a list of synonymous words or collocations.

**Consequence:** a concept cannot exist without an English term to denote it

e.g. no synset for fr:fleuve (large river that flows into a see or ocean)

WordNet is not an ontology creation/management system

## Some synsets with the word "plan"

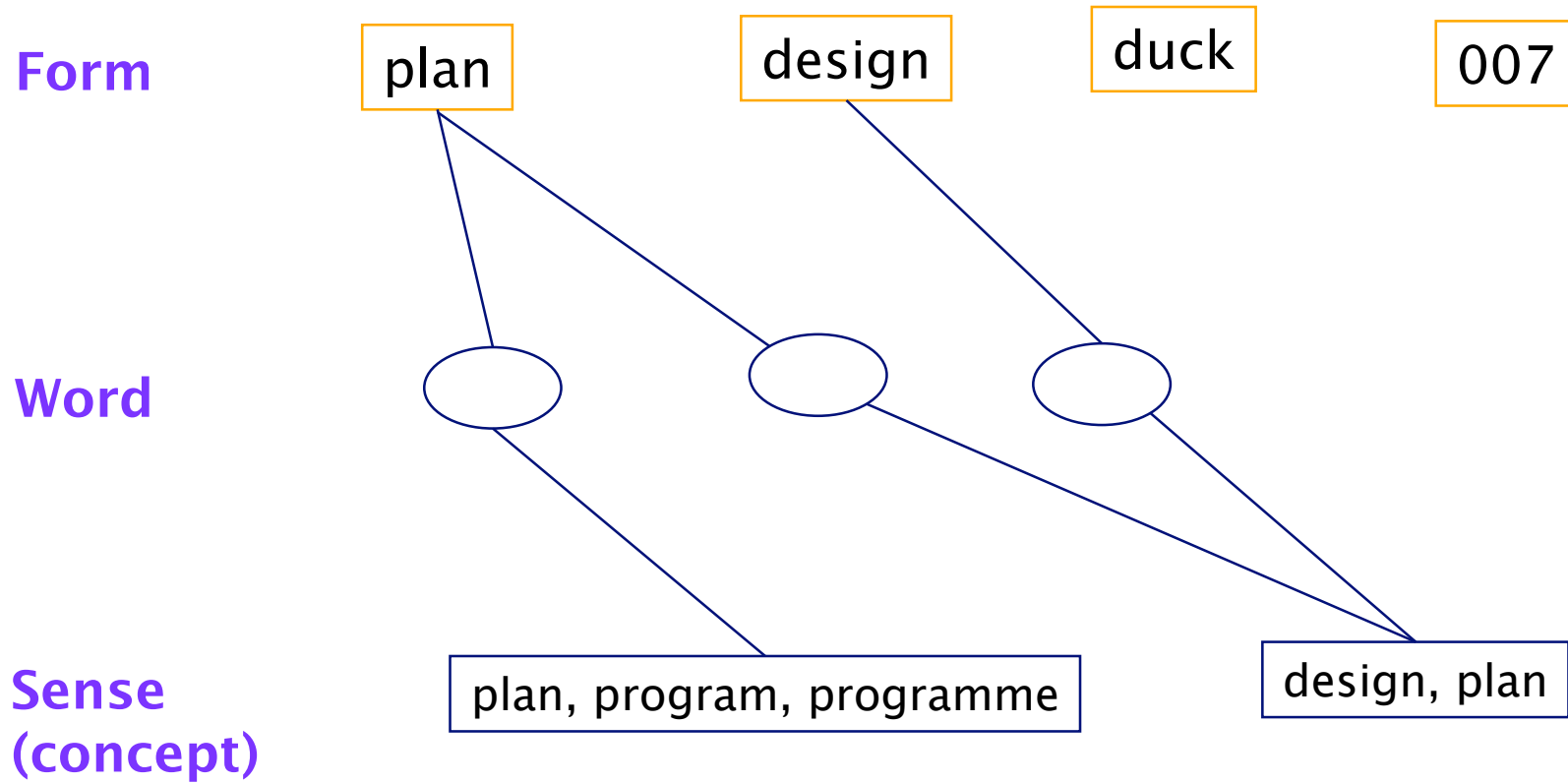
### Noun

- \* S: (n) **plan**, **program**, **programme** (a series of steps to be carried out or goals to be accomplished) "they drew up a six-step plan"; "they discussed plans for a new bond issue"
- \* S: (n) **design**, **plan** (an arrangement scheme) "the awkward design of the keyboard made operation difficult"; "it was an excellent design for living"; "a plan for seating guests"
- \* S: (n) **plan**, **architectural plan** (scale drawing of a structure) "the plans for City Hall were on file"

### Verb

- \* S: (v) **plan**, **be after** (have the will and intention to carry out some action) "He plans to be in graduate school next year"; "The rebels had planned turmoil and confusion"
- \* S: (v) **plan**, **project**, **contrive**, **design** (make or work out a plan for; devise) "They contrived to murder their boss"; "design a new sales strategy"; "plan an attack"

# Forms, words and senses



## Lexical and semantic relations

Lexical relations hold between semantically related word forms

**S:** (n) **plan, program, programme** (a series of steps to be carried out or goals to be accomplished) "they drew up a six-step plan"; "they discussed plans for a new bond issue"

- direct hyponym / full hyponym
- direct hypernym / inherited hypernym / sister term
- derivationally related form
  - **W:** (v) **program** [Related to: program] (arrange a program of or for) "program the 80th birthday party"
  - **W:** (v) **programme** [Related to: programme] (arrange a program of or for) "program the 80th birthday party"

# Semantic relations

Hold between synsets (senses)

Nouns

- hyponym/hypernym
- instance
- meronym/holonym

"is a"

"is an instance of"

"has part"

Verbes

- troponym/hypernym
- antonym
- cause
- etc.

"a way of"

Adjectives

- antonym
- etc.

# Hyponym/hypernym and instance

## Relations between nouns

X has a hyponym Y if the sentence ***Y is a X*** is true.

<i>a cat is an animal</i>	<i>subclass</i>
<i>a state capital is a city</i>	<i>subclass</i>
<i>Belfast is a city</i>	<i>instance</i>

Corresponds either to the **subclass** or the **instance** relations in formal ontologies.

In WN: instance  $\subseteq$  hyponym

$a -(\textit{instance})-\> b \Rightarrow a -(\textit{hyponym})-\> b$

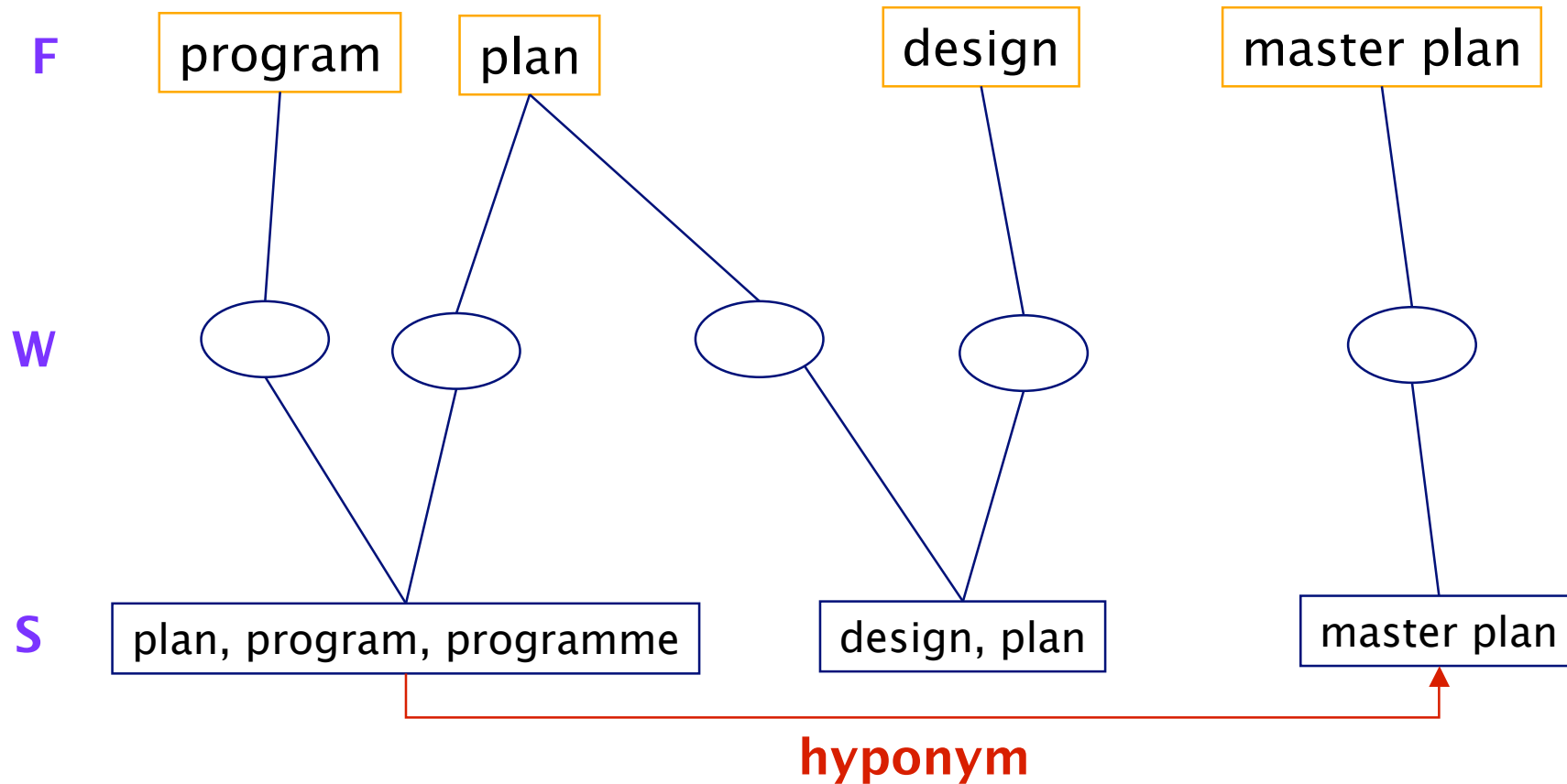
## Example

S: (n) **plan, program, programme** (a series of steps to be carried out or goals to be accomplished) "they drew up a six-step plan"; "they discussed plans for a new bond issue"

\* **direct hyponym / full hyponym**

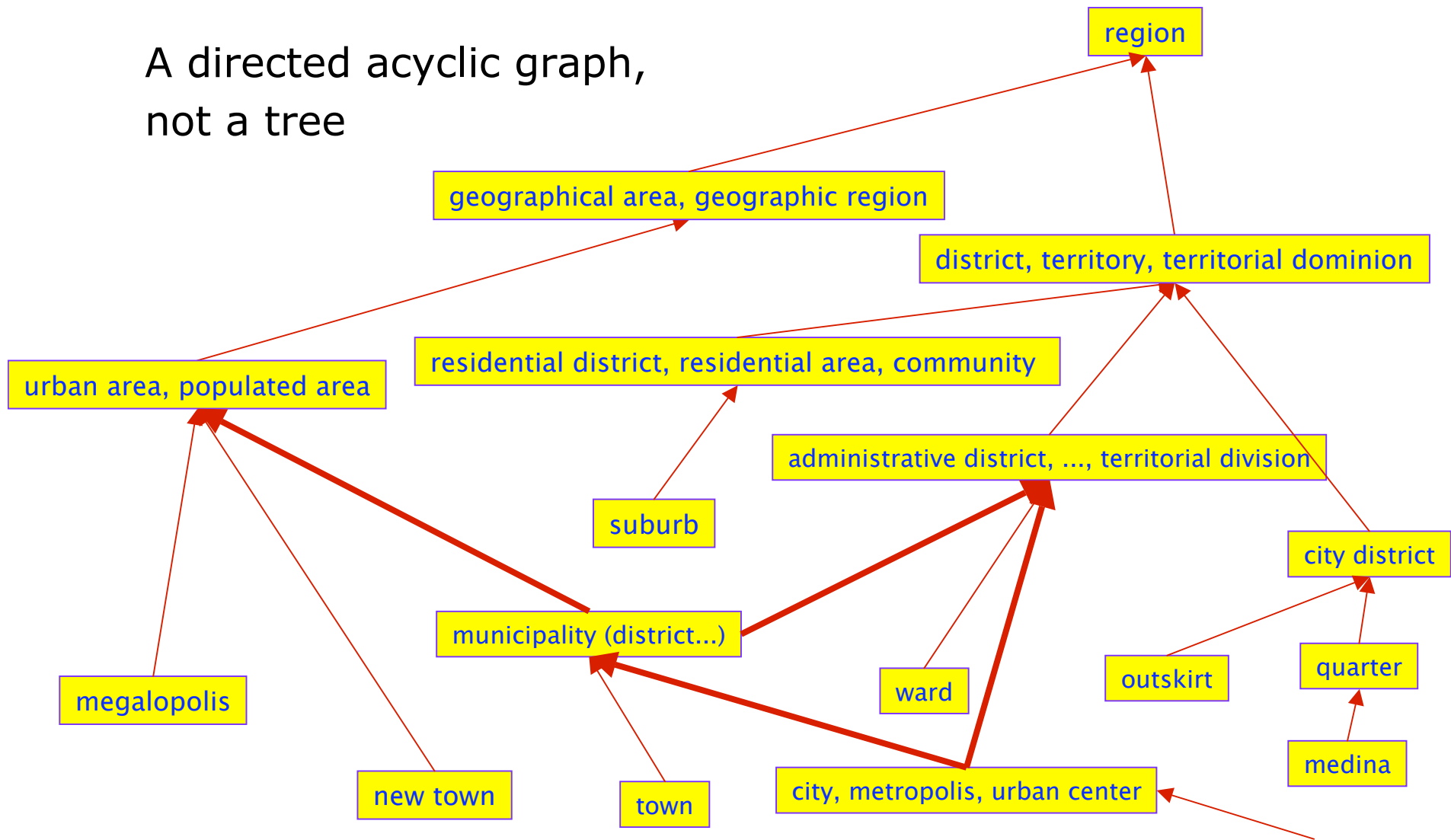
- S: (n) **audit program, audit programme** (a listing of audit procedures to be performed in completing an audit)
- S: (n) **outline, schema, scheme** (a schematic or preliminary plan)
- S: (n) **master plan** (a long-term outline of a project or government function) "the zoning board adopted a master plan for the new development"
- S: (n) **blueprint, design, pattern** (something intended as a guide for making something else) "a blueprint for a house"; "a pattern for a skirt"

# Position of the semantic relations



# Hypernym Hierarchy

A directed acyclic graph,  
not a tree



## Meronym/holonym (between nouns)

The meronym relation holds between X and Y if Y is a part of X or if X is made of Y (substance).

<i>a financial center is a part of a city</i>	<i>part</i>
<i>bronze is made of copper</i>	<i>substance</i>
<i>a street is made of a pavement</i>	<i>substance</i>
<i>a pavement is made of paving stones</i>	<i>substance</i>

## The Size of WordNet

≥ 166 000 words = pairs (form, sense)

118 000 forms

90 000 senses

polysemic words: 17%

words with a synonym: 40%

# WordNet and UCE

Preliminary study

Comparison with Urbanet

- look for concepts of UN in WN
- look for concepts of WN in UN

## Urbanet thesaurus

Ministère de l'Équipement, des Transports et du Logement  
(France)

*The town planning site of the Ministry of Public Works, Transport  
and Housing GUHC Directorate General for Town Planning,  
Housing and Building I*

approx. 4000 entries

technical realisation : CETE Nord-Picardie

# Organisation of Urbanet

Hierarchy of concepts and topics

Hierarchical relations have several meanings

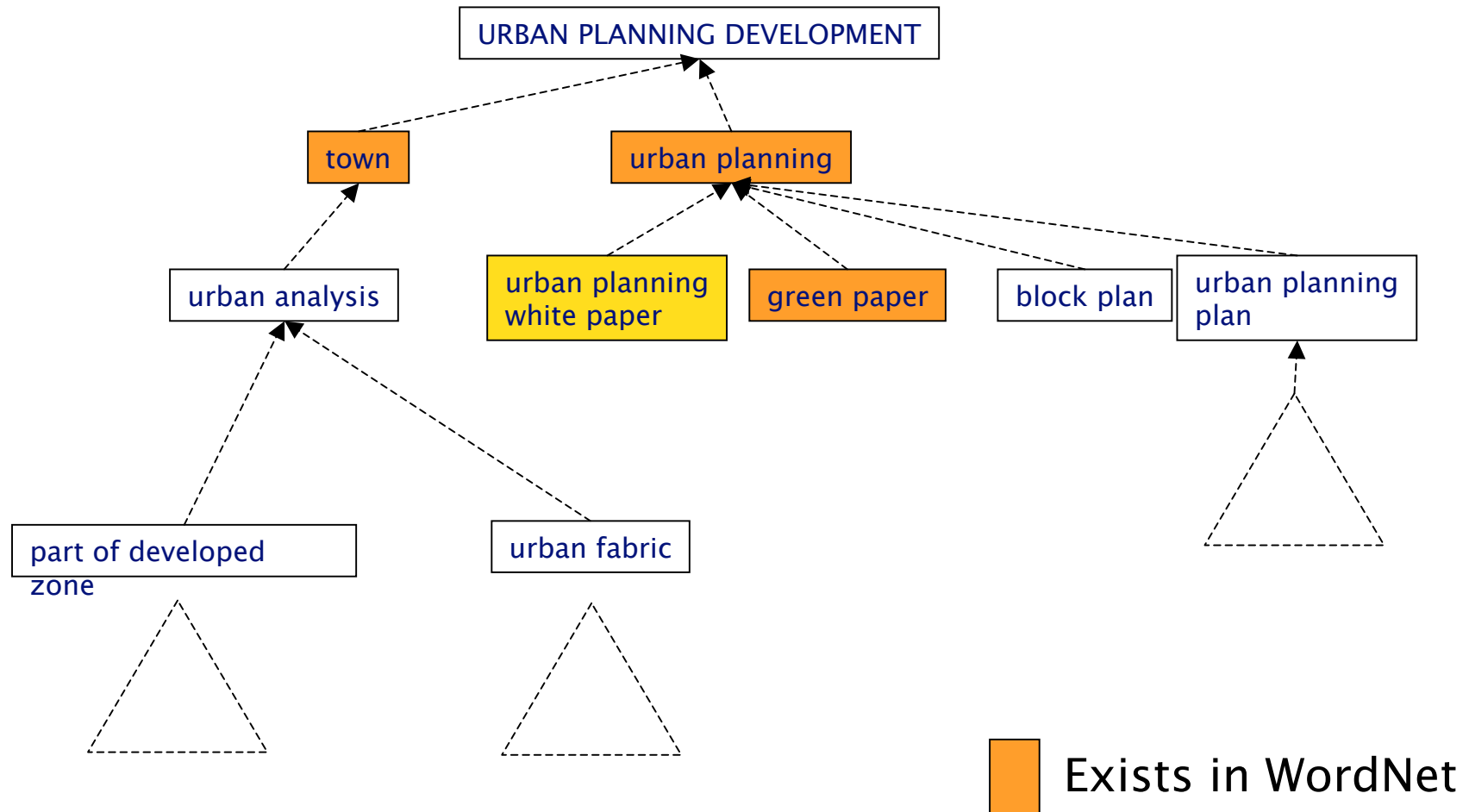
- "subtopic"

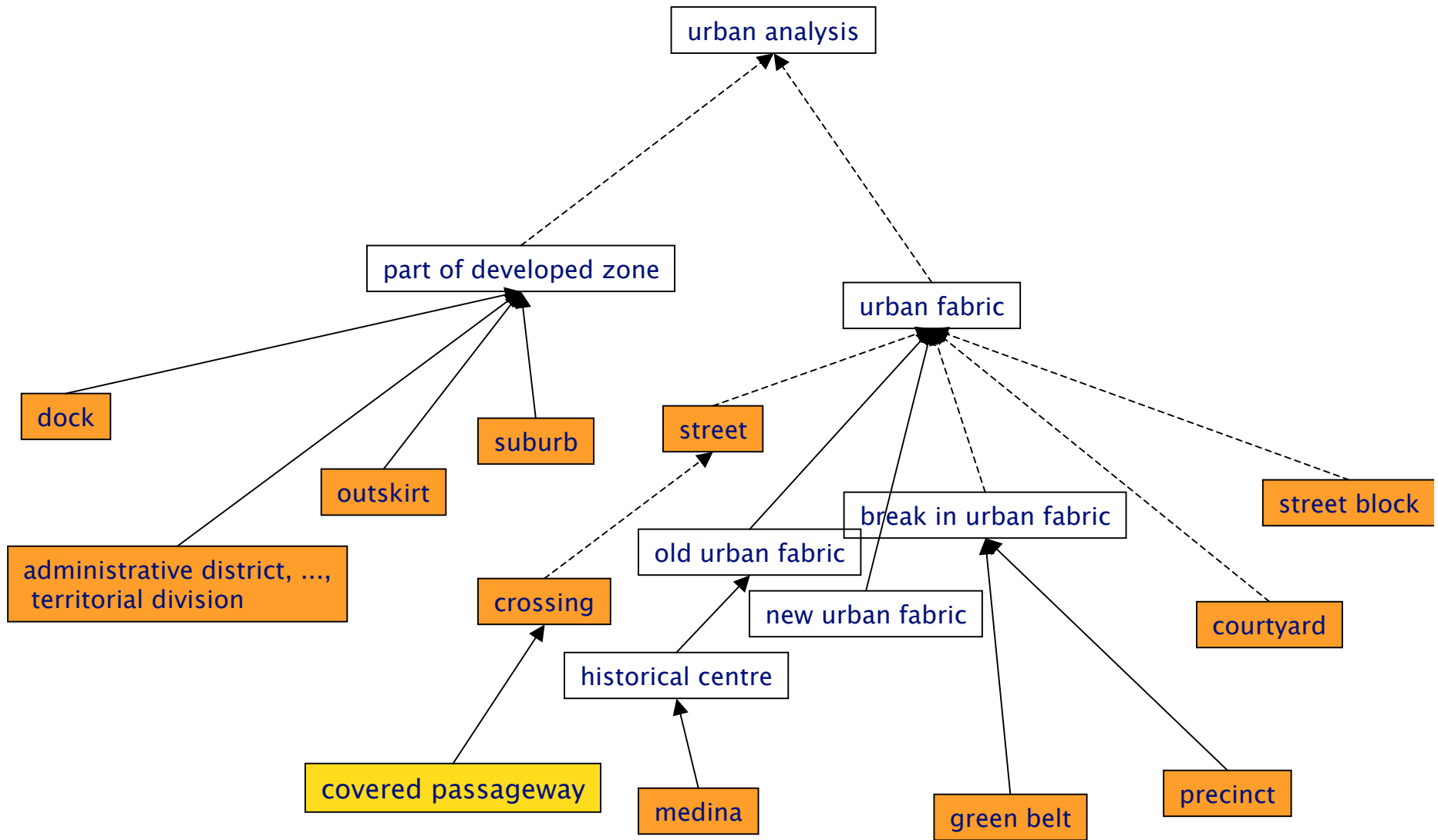
- "is a"

- "is a part of"

- etc.

# Urbanet Structure







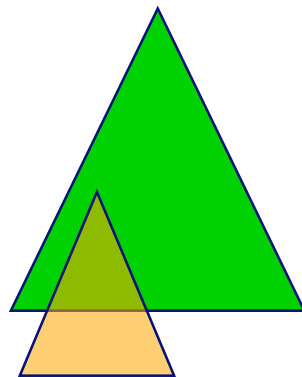
## Observations

- Urbanet concepts do not form a subtree or subgraph of WN
- They are scattered across WN
- WN and Urbanet represent **different worldviews**

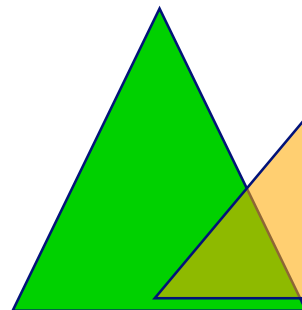
## Observations

(to be confirmed)

- Many specific UN concepts appear to WN
- Several abstract UN concepts do not appear in WN

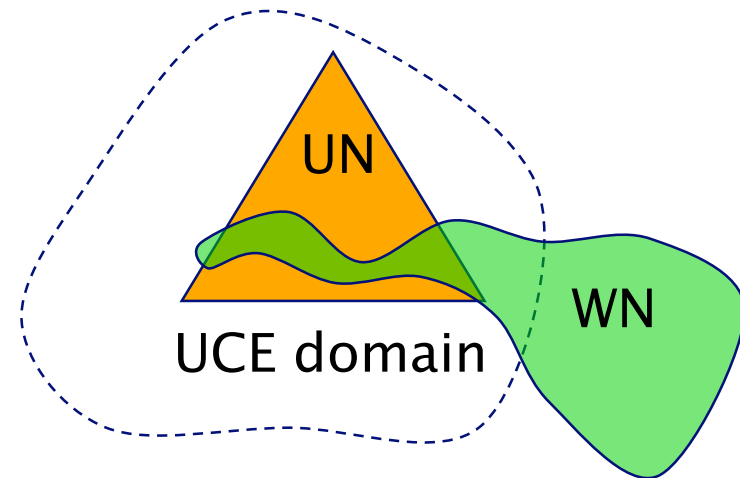
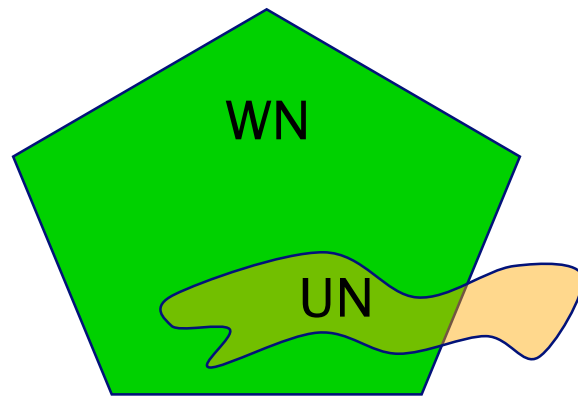


not like that



more like this

# Intersections



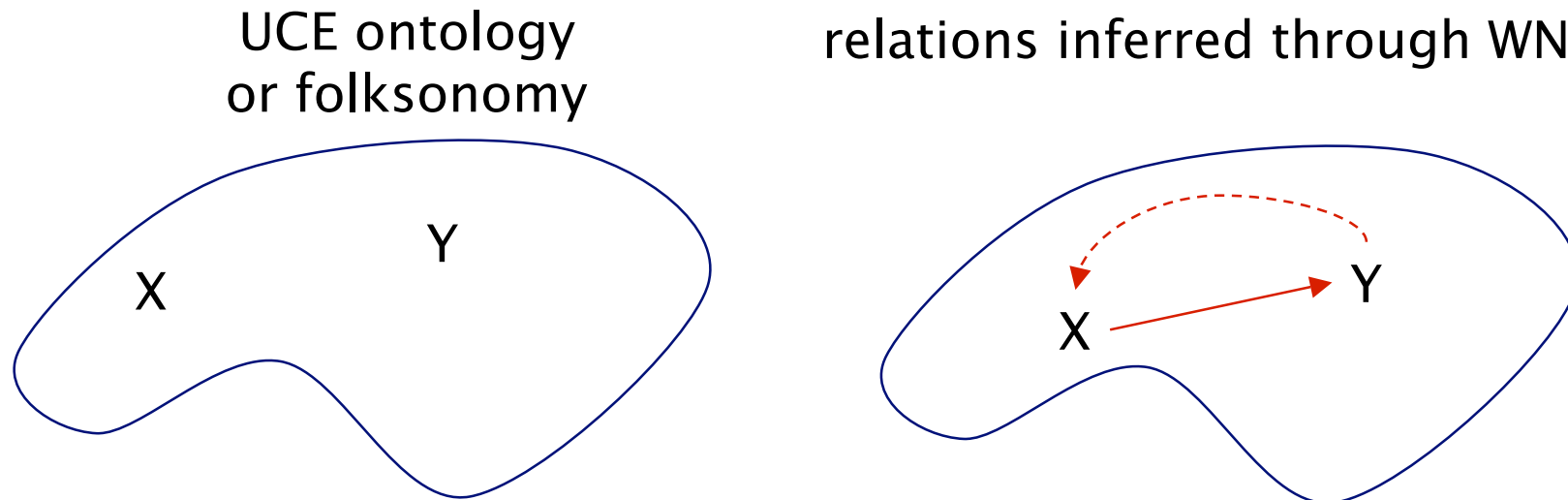
## Possible uses of WN in UCE applications

Adding/infering structure (is-a, part-of)

WN has well defined relations

Can be used to re-classify UCE concepts

= project a UCE ontology onto another worldview



## Possible uses of WN in UCE applications (2)

### User interface of UCE applications / UCE resources

- simple linguistic processing (synonyms)
  - natural/general language access to UCE concepts
    - designating UCE concepts with everyday words
    - access for non-specialists / public access
  - WN concept browser to access UCE concepts
    - the WN hierarchy maybe more intuitive/usual
  - multilingual interfaces (through EuroWordNet)
- => Necessary to **align** UCE ontologies with WN

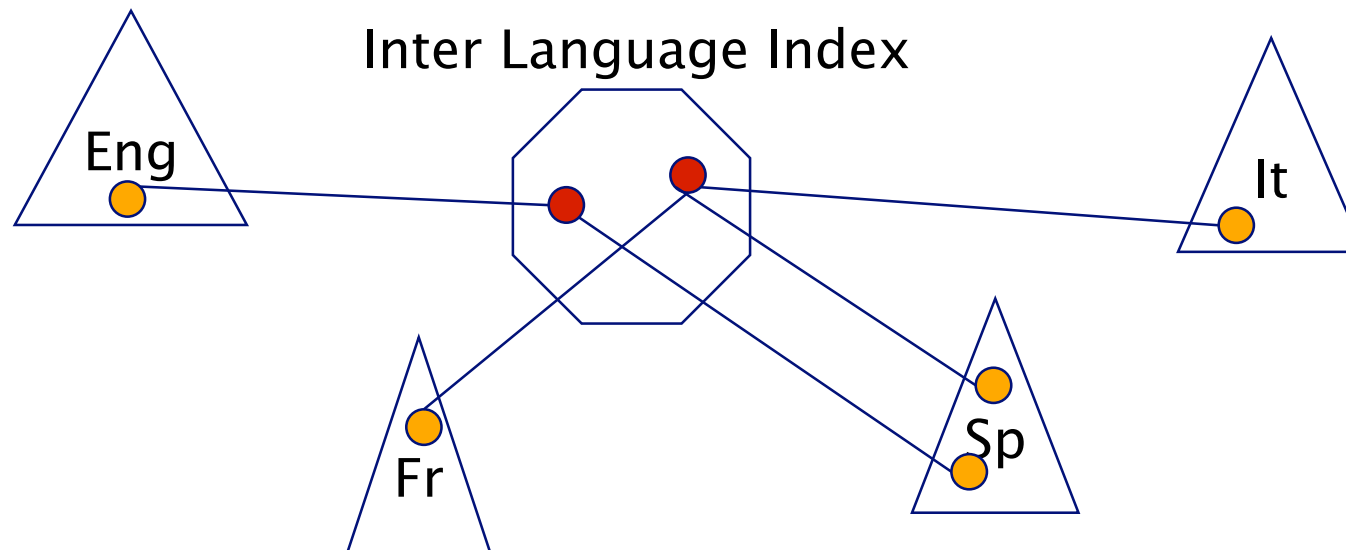
# Alignment

Building bridges between ontologies

Automatically or semi-automatically

Active research field (related to the semantic Web idea)

Example: EuroWordNet = set of aligned WordNets



## Conclusion

WordNet is not a UCE ontology

It does not contain UCE

Possible uses

- adding information/structure to UCE ontologies or folksonomies
- at the user interface level
  - need alignment
- others ....