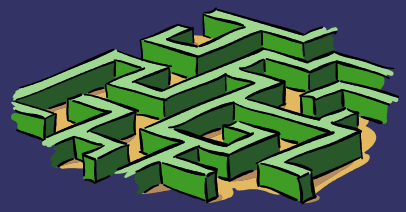
# *Universal Dictionary of Concepts*

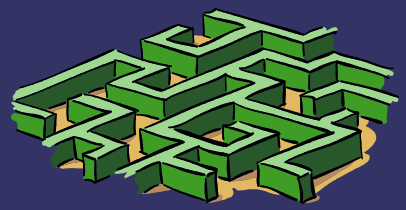
***Dictionary of the semantic pivot language UNL and its network of concepts***

Boguslavsky I.M.
Dikonov V.G.
IITP RAS



# *Goals of the dictionary*

- The UNL project unites efforts of several groups of linguists and information scientists from Russia, Spain, France, India, Egypt and other countries.
- **Each group is responsible for the support of one or several national languages** and develops the necessary software and relevant linguistic resources.
- Different groups evolved their own dialects of the artificial language with different dictionaries of concepts, not quite compatible with each other.
- The new resource aims to unite and replace multiple existing UNL dictionaries and become **the common standard lexicon of the UNL language**.



# *Universal Words*

- The creators of the UW format made a lot of effort to make the words of UNL self-explanatory.

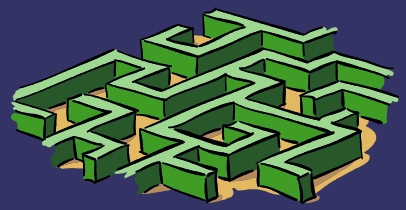
  ***headword (icl>hypernym>class, equ>synonym, agt>class, obj>class)***

- Each UW consists of a headword and a list of restrictions, which are used to narrow the semantics of the headword and remove its ambiguity.
- There are three types of restrictions:
  - **Ontological** codify the general knowledge about the world: **icl** (inclusion into a class), **pof** (part of), **iof** (instance of).

    *tongue(**icl**>concrete_thing,**pof**>body)*
  - **Semantic** help to distinguish between concepts that have one common headword: **equ** (equivalent), **ant** (antonym), **com** (component), **fld** (domain).

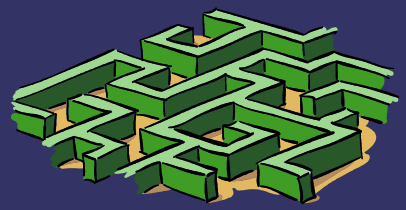
    *ably(icl>how,**equ**>competently,**ant**>incompetently,**com**>skill)*
  - **Argument** reflect the typical argument frame in terms of UNL relations: **agt** (agent), **cag** (co-agent), **obj** (object), **plc** (place), **tim** (time), **rsn** (reason)...

    *buy(icl>get>do,**agt**>person,**obj**>thing,**cob**>thing,**src**>thing)*
- Each UW should have only the **minimal** set of restrictions necessary to express the difference between concepts with UWs sharing the same headword.



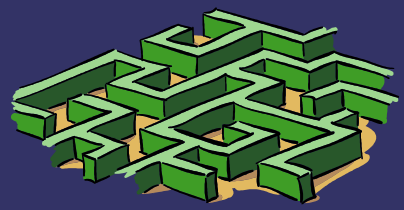
# *Features of the Dictionary*

- The greatest difference between the dictionary of UNL and lexicons of natural languages is that **its units are not polysemous**.
  - For technical reasons the dictionary may contain several UWs representing the same concept.
- **The dictionary provides ontology-like information** in the form of links relating the concepts with each other and putting them into different semantic classes.
- Morphological and syntactic features of the natural language words are not relevant for UNL.
- Being the lexicon of a pivot language it must **equally well reflect the wealth of concepts** existing in **all languages.**
  - The dictionary itself can be used as a pivot to translate words. If there is no direct translation, the links between concepts help to find the best possible equivalent in any of the supported languages.



# *The UNL Dictionary and Wordnet*

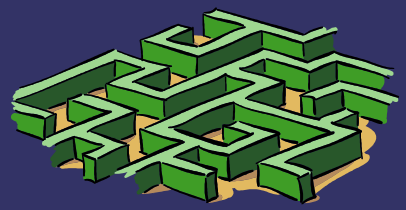
- The Universal Dictionary of Concepts has much in common with Wordnet.
- Princeton Wordnet was used as the source for generation of most currently existing UWs and the English local dictionary. The **concepts derived from Wordnet have back references to PWN v2.1.**
- Differences between the new dictionary and Wordnet:
  - No bias towards any single natural language..
  - **Members of Wordnet synsets are treated as quasi-synonyms**, which may have subtle differences.
  - Different organization of the semantic network (including **polyhierarchy** instead of tree structure, argument structure).
  - UNL dictionary provides generalized **semantic frame** description.
  - No separation between parts of speech.
  - **Includes prepositions and conjunctions**.
  - Concepts found to be missing from Wordnet are added.
  - English words not described by Princeton Wordnet, e.g. some phrasal verbs, are added as well as any non-English concepts.



# *Structure of the Dictionary*

- Universal Dictionary of concepts consists of three parts:

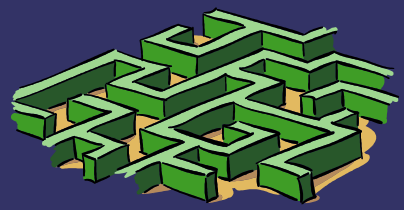
1. **List of concepts**
   defines the inventory of UWs.
2. **Local dictionaries**
   link concepts with words and expressions of natural languages.
3. **Semantic Network**
   establishes relations between concepts.



# List of Concepts

- The list of UWs includes all concepts existing in the dictionary and the UNL language.
- There is no discrimination between UWs based on the source natural language. New UWs may be created on the basis of UWs representing concepts from any language,

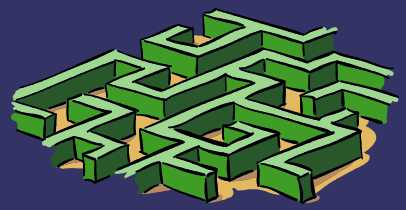
| UW | Legacy forms | Exists in | Added by | Wordnet |
|---|---|---|---|---|
| man(icl>person,ant>animal,equ>human) | man(icl>person) | ru, en... | Vyacheslav | |
| person(icl>abstract_thing,equ>personality) | | ru, en... | Vyacheslav | |
| one(icl>person) | | ru, en... | Vyacheslav | |
| mankind(icl>homo>thing,equ>world) | | ru, en... | Spanish center | mankind %1:05:00::6 |
| human(icl>hominid>thing,equ>homo) | | ru, en... | Spanish center | human %1:05:00::3 |
| between(icl>how,obj>thing,plc<uw) | | ru, en... | Tatiana | |
| between(icl>how,com>quantity,obj>thing) | | en... | Tatiana | |
| between(icl>how,com>participation,obj>thing) | | ru, en... | Tatiana | |



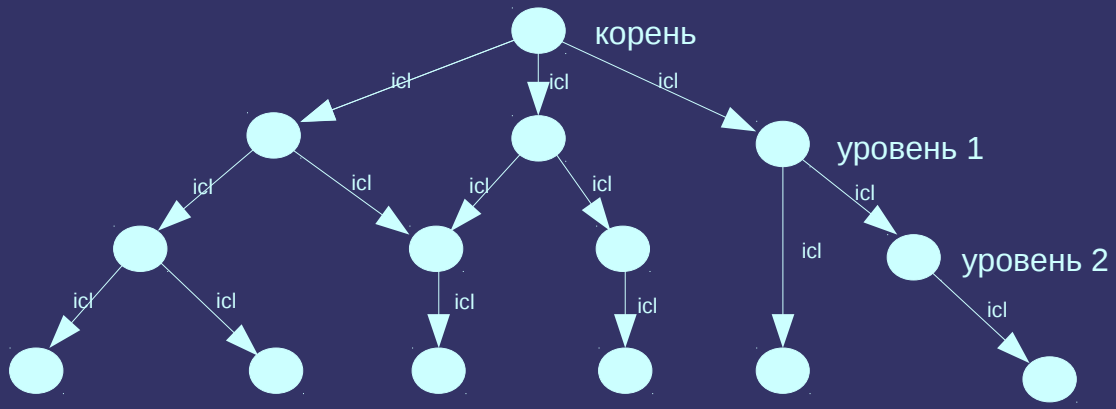
# *Semantic Network*

- The concepts will be organized into a semantic network linked by the relations of synonymy, antonymy, hypernymy, association and various argument frame relations.
- Tracing the relations between concepts helps to find the nearest equivalent if there is no direct translation of a concept into the required language.

  For example: In order to translate into English the Russian word “*жениться*”, correctly rendered in UNL with the concept “*to acquire a wife*”, the system must replace this concept with its hypernym “*to marry*”, because there is no English word with exactly the same meaning.
- The proposed semantic network consists of three superimposed structures made of different types of relations.



# *Semantic Network*

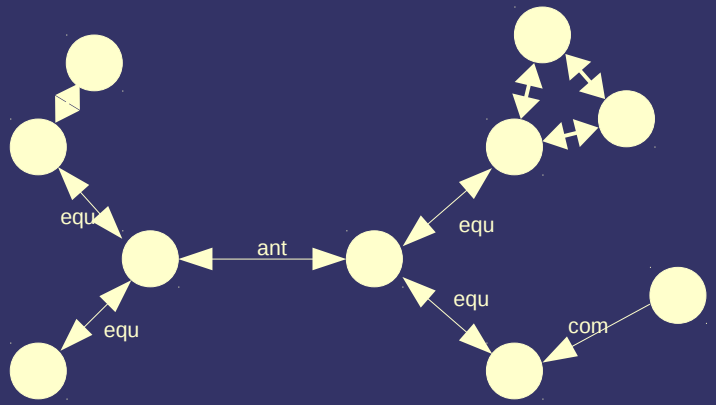


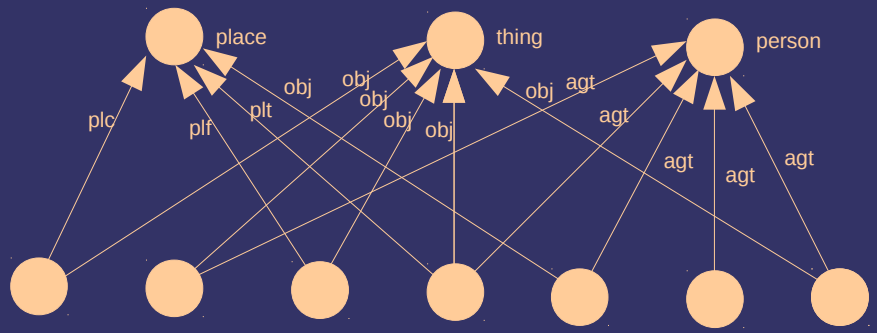

«Ontological» relations: hypernymy, meronymy, instantiation

Unlike Wordnet the UNL dictionary permits polyhierarchy and facet classification of concepts.

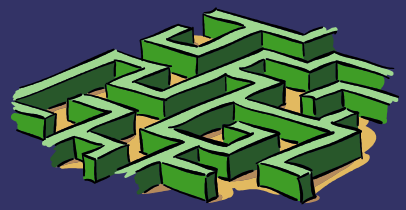
«Semantic» relations: synonymy, antonymy, association, domain

This structure may consist of many isolated fragments.

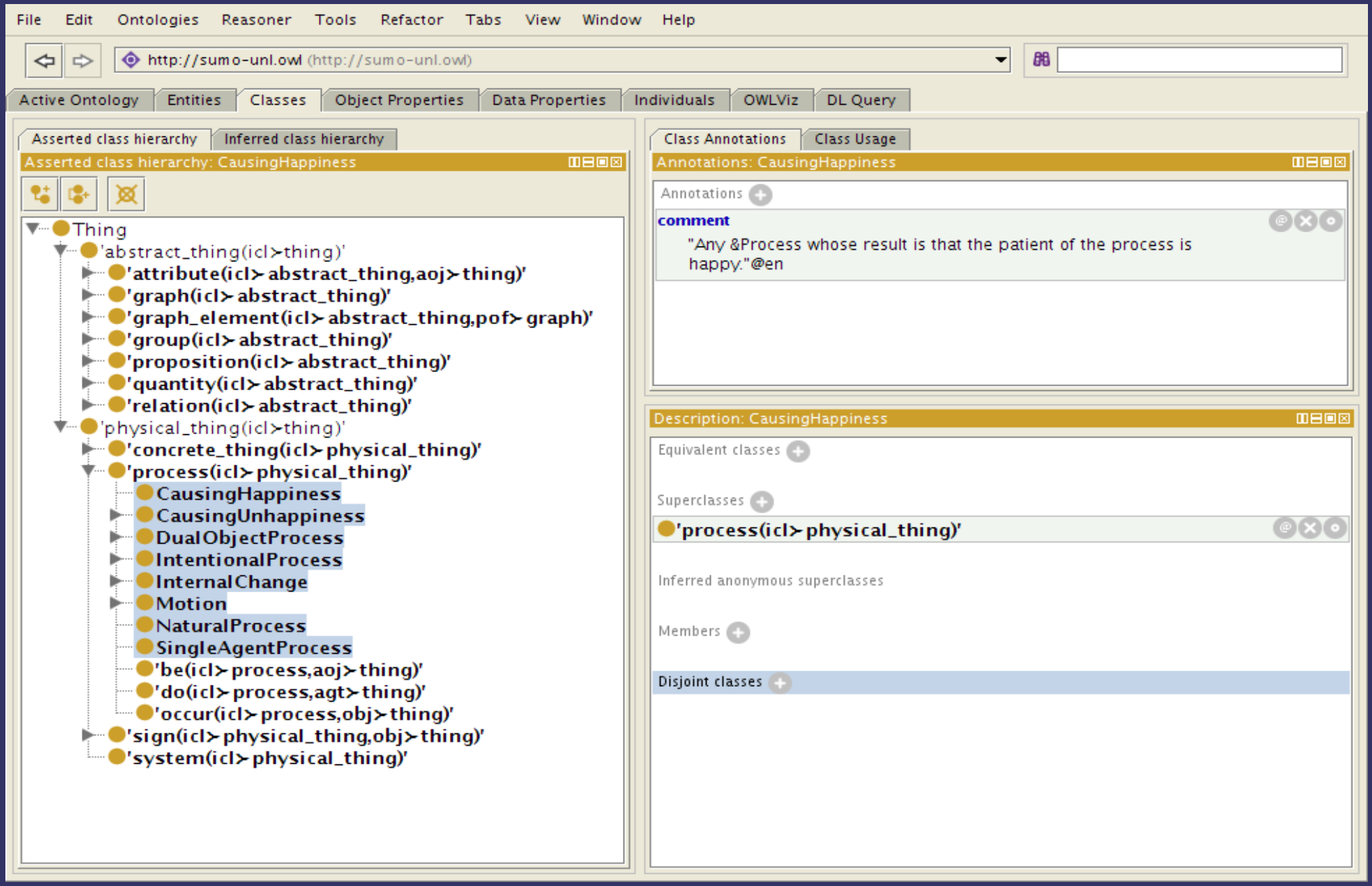



«Argument» relations



# Merging with Ontology



# Local Dictionaries

- Provide translations of UWs into the corresponding natural language.
- May contain additional (public or private) data fields, e.g. definitions and examples or any other information useful for generation of text in the natural language.
- Translations and all public data provided by the local dictionary authors will be an integral part of the dictionary.

| UW | ENG Word | POS | ENG Comment | ENG Example |
|---|---|---|---|---|
| man(icl>person,ant>animal,equ>human) | man | N | any human being | group of three people |
| person(icl>abstract_thing,equ>personality) | person | N | the personality of a human being | a nice person |
| one(icl>person) | one | N | any person as representing people in general | one should never be complacent |
| mankind(icl>homo>thing,equ>world) | mankind, man, world | N | all of the living human inhabitants of the earth | one giant leap for mankind |
| reformed_gas(icl>matter,fld>chemistry fld>engine) | reformed gas | A←N | gas mixture produced by pyrolysis | conversion of fuel into a reformed gas |
| reformer(icl>converter>functional_thing, fld>chemistry,fld>engine) | reformer | N | a chemical reactor | catalytic steam reformer |

| UW | RUS Word | POS | RUS Comment | RUS Example |
|---|---|---|---|---|
| man(icl>person,ant>animal,equ>human) | человек | N | человеческое существо | отряд в сорок человек |
| person(icl>abstract_thing,equ>personality) | человек | N | личность, персона | приятный человек |
| one(icl>person) | человек | N | всякий, любой | человек не должен себя ронять |
| mankind(icl>homo>thing,equ>world) | человек | N | человеческая цивилизация | человек шагнул в космос |
| reformed_gas(icl>matter,fld>chemistry, fld>engine) | синтез-газ | N | газовая смесь, результат пиролиза | конверсия керосина в синтез-газ |
| reformer(icl>converter>functional_thing, fld>chemistry, fld>engine) | термокаталитический конвертер | A←N | химический реактор | газовую смесь получают на борту в термокаталитическом конвертере |

# UW Construction Wizard

Word: drive  Source language: en

Comment: transport in a vehicle

Example: to drive smb. (obj) to the station (plt)

English Comment: transport in a vehicle

Already existing UWs:

| | |
|---|---|
| drive(icl>be,aoj>thing) | have certain properties when operated / my new truck drives well |
| drive(icl>carry>do,agt>person,ins>vehicle,obj>thing,plf>thing,plt>thing) | transport in a vehicle / to drive smb. (obj) to the station (plt) |
| drive(icl>device>thing) | (computer science) a device that writes data onto or reads data from a storage medium / |
| drive(icl>force>do,gol>uw,obj>thing) | force into a state or to do smth / to drive smb. (obj) to steal (gol) |
| drive(icl>golf_stroke>thing) | hitting a golf ball off of a tee with a driver / he sliced his drive out of bounds |
| drive(icl>journey>thing) | a journey in a vehicle (usually an automobile) / he took the family for a drive in his new car |

drive(icl>carry>do,agt>person,ins>vehicle,obj>thing,plf>thing,plt>thing)

1. Headword:

drive

2. Ontological restrictions:

Top category: Verbal action  Closest hypernym: carry  Instance of:  Is a part of:

3. Semantic restrictions:

Synonym:  Antonym:  Important component of definition:

Subject domain:  Is a typical modifier of:

4. Semantic argument frame:

| Argument role | Most general suitable class to fill in the slot | Reverse | |
|---|---|---|---|
| Agent | Thing → Concrete → Living → Human bein | ☐ | A participant that initiates an action. Normally it is a living being, but non-living causes of events also belong here. |
| Instrument | vehicle | ☐ | Defines a material object which serves as an instrument for doing something. |
| Object | Any Thing | ☐ | A participant directly affected by the action. Besides, this relation is used in a wide range of situations to connect the most important participant distinct from the agent. |
| Place from | Any Thing | ☐ | Initial point of movement, maybe imaginary. |
| Place to | Any Thing | ☐ | Final point of movement, maybe imaginary. |

Such Universal Word already exists in the dictionary, Please, add more restrictions.

Reset  Save UW  Exit