

Lexical Ontology of Persian Verbs

Artem Lukanin (ice_lc@mail.ru), South Ural State University

Persian Verb Conjugator with Integrated Persian WordNet

<http://persian.nmelrc.org/pvc/>

- a linguistic network of frames, which maintains the syntagmatic relationships of Persian verb morphemes and paradigmatic relationships between word-form frames,
- together with the algorithms of its processing, is used to generate Persian verb forms from the infinitives;
- initially the lexicon had only exceptions to the rules of the present stem derivation [3];
- later it has also become useful as a Persian-English thesaurus of verbs:
 - examples of verb usage;
 - a set of relationships between verbs, that maintain styles and registers, semantics and syntax of the Persian language.

MAIN RELATIONSHIPS

Synonymy

- the synonymous meanings of the verbs are grouped into synsets (sets of synonyms, [4]);
- to help the lexicographer populate the database we developed a special user interface: for each Persian verb they can add several meanings (English translations) and include every meaning into a synset.

Transcription		
Infinitive	Present Stem	
<input type="text" value="nafas keshidan"/>	<input type="text" value="nafas kesh"/>	
Note: aa=ā		
Meaning	Tags	add meaning
<input type="text" value="to breathe"/>	<input type="text" value="n/a"/>	select a synset
<i>draw air into, and expel out of, the lungs</i> [remove from synset]		
<input type="text" value="نفس کشیدن to breathe"/>	<input type="text" value="دمیدن to breathe"/>	<input type="text" value="نفس زدن to breathe"/>
<input type="text" value="دم زدن to breathe"/>	<input type="text" value="استنشاق کردن to breathe"/>	<input type="text" value="تنفس کردن to breathe"/>
<input type="text" value="to inhale"/>	<input type="text" value="n/a"/>	select a synset
<i>draw in (air)</i> [remove from synset]		
<input type="text" value="نفس کشیدن to inhale"/>	<input type="text" value="نفس داخل دادن to inhale"/>	<input type="text" value="در دمیدن to inhale"/>
<input type="text" value="هوا فرو بردن to inhale air"/>	<input type="text" value="استنشاق کردن to inhale"/>	

Synset Relationships

Integrated into the verb definitions is PVC's WordNet, based as closely as possible on Princeton English WordNet 3.1. We use four relation types, presented in [1, p. 54, Figure 3]:

- troponymy (hypernym ↔ troponym), e.g. *نفس کشیدن* /nafas keshidan/ "to breathe" (draw air into, and expel out of, the lungs) ↔ *خرخر کردن* /khor-khor kardan/ "to snore" (breathe noisily during one's sleep);
- entailment, e.g. *خرخر کردن* /khor-khor kardan/ "to snore" (breathe noisily during one's sleep) → *خوابیدن* /khābidan/ "to sleep" (be asleep);

- backward presupposition, e.g. *باختن* /bākhtan/ to lose (fail to win) → *مسابقه دادن* /mosābeqe dādan/ "to compete" (compete for something; engage in a contest; measure oneself against others);
- cause, e.g. *جان گرفتن* /jān gereftan/ "to take a life" (cause to die; put to death, usually intentionally or knowingly) → *مردن* /mordan/ "to die" (pass from physical life and lose all bodily attributes and functions necessary to sustain life).

However, Princeton WordNet does not single out backward presupposition, and unites it with entailment.

to breathe (draw air into, and expel out of, the lungs)

Syn: *دمیدن, نفس زدن, دم زدن, استنشاق کردن, تنفس کردن*

entailment

- *draw in (air):* *نفس کشیدن, نفس داخل دادن, در دمیدن, هوا فرو بردن, استنشاق کردن*
- *expel air:* *نفس بیرون دادن, باز دمیدن, دم بر آوردن*

troponymy

- *breathe noisily during one's sleep:* *خرخر کردن*
- *heave or utter a sigh; breathe deeply and heavily:* *اه کشیدن, دم بر آوردن*
- *breathe noisily, as when one is exhausted:* *نفس نفس زدن*
- *breathe with difficulty:* *خس خس کردن*

Different teams already make wordnets for the Persian language [2; 5], however our resource has some advantages:

- the results of our work can be seen immediately on the PVC site, which actually comprises different types of resources at the same time;
- it is a paradigm generator for any Persian verb and a Persian-English thesaurus;
- the dictionary database and the generation algorithms are used for generation of tests on the Persian language.

ACKNOWLEDGMENTS

The main lexicographer of PVC is Constance Bobroff. The names of the top 10 sentence contributors are listed at <http://persian.nmelrc.org/pvc/sentences.php?contributor>.

REFERENCES

- [1] Fellbaum, C. "English Verbs as a Semantic Net", *5 papers on WordNet*. [Online], Revised August 1993, Retrieved (July, 2, 2012), p. 40-61. Available: <http://wordnetcode.princeton.edu/5papers.pdf>
- [2] Keyvan, F., Borjian, H., Kasheff, M., Fellbaum, C. Developing PersiaNet: The Persian Wordnet. In Proceedings of the 3rd Global WordNet conference, South Korea. 2006. Pp. 315-318
- [3] Lukanin, A., Bobroff, C. Frame approach to Persian verb generation for educational purposes. In the Proceedings of the 2nd Workshop on Computational Approaches to Arabic Script-based Languages. — Linguistic Institute, Stanford, California, USA. July 21-22, 2007. Pp. 98-105.
- [4] Miller, George A. WordNet: A Lexical Database for English. In Communications of the ACM, 1995. 38 (11). Pp. 39-41.
- [5] Rouhizadeh, M., Yarmohammadi M. A., Shamsfard, M. Developing The Persian WordNet Of Verbs: Issues Of Compound Verbs And Building The Editor. In the Proceedings of The 5th International Conference of the Global WordNet Association (GWC-2010). Mumbai, India, 2010.