

# An Ontology for a Section of the *Ashtadhyayi*: A Final Report

Alok Debnath

20161122

alok.debnath@research.iiit.ac.in

---

## Abstract

In this report, I demonstrate the development of an ontology for a subset of the words present in Panini's *Ashtadhyayi*. The set of words provided were in fact not related to one another, rather were in lexicographic ordering. This handicap made the creation of this ontology an interesting exercise. I did not use existing knowledge graphs, rather used a simple approach of an abstract notion of semantic role of the words. This allows for a simple and flexible ontology, but it also has the drawbacks of not representing polysemy and some corner cases based on the definitions of each category.

*Keywords:* Sanskrit, Ontology, Ashtadhyayi

---

## 1. Introduction

Natural language ontology is a branch of both metaphysics and linguistic semantics. Its aim is to uncover the ontological categories, notions, and structures that are implicit in the use of natural language, that is, the ontology that a speaker accepts when using a language [1]. Sanskrit has a very rich tradition of arrangement and representation of meanings, as has been demonstrated in the *Nyaya* and *Vaisheshika* schools.

More recently, the work done in knowledge representation in natural language processing has led to the conception and development of knowledge graphs and networks [2]. While a consequence of semantic representation of knowledge, these are not strictly ontological in nature [3], as they do not show the hierarchical classification of the meanings of words based on their use and intrinsic properties. Rather, these representations are arrangements of words and the relations between those words as one might find them in a

lexicon. Some examples of these types of representations include WordNet [4], VerbNet [5] and ConceptNet [6].

In this project, our aim was to select a set of about a hundred Painian semantic conditions used in either the Dhatupaha or in the Ahadhyayi and construct a formal semantic network, investigate various existing formal semantic networks and select one to use as the model or framework for your network of Painian semantic conditions. The semantic conditions chosen were from the Astahayi. I investigated the Vaisheshika ontology (categories of *padartha*; section 2) and the network WordNet (section 3). However, given that the words in the subset I chose were lexicographically chosen, it would be futile to use a description as large as the Vaisheshika ontology. On the other hand, the words were not related enough to use a lexical arrangement such as WordNet, nor were they conceptually intertwined as the ConceptNet, which is why I chose to restructure a subset of the Vaisheshika ontology to suit the words I chose. This process has been described in Section 4

## 2. Vaisheshika ontology

The most basic concept in the Vaisheshika ontology is that of *padartha*. The word *padartha* is a sandhi of the words *pada*, which means word and *arhta* which translates to meaning. Originally, the word *pada* implies that form of a word which can be used in a sentence, including its inflectional changes. *Padartha* therefore means "the meaning or referent of words" [7].

*Padartha* is the basic unit of the ontology of both, the Nyaya and the Vaisheshika ontologies. The original Vaisheshika ontology categorizes the *padarthas* into six categories [8]:

- *dravya* (substance): This semantic category is defined as those concepts which are represented as objects or substances. In these substances, there are those which can be perceived by one or a combination of senses, as well as the objects of intellectual or logical existence, with no "real" or perceivable manifestation.
- *gua* (quality): A *guna* is a quality or a property of substances. A *guna* is not capable of independent existence, as its function is only of describing the properties or qualities of substances, activities and other *gunas*. There are 17 types of *gunas* in the original *Vaisheshika sutra*, while later authors have added 7 more.

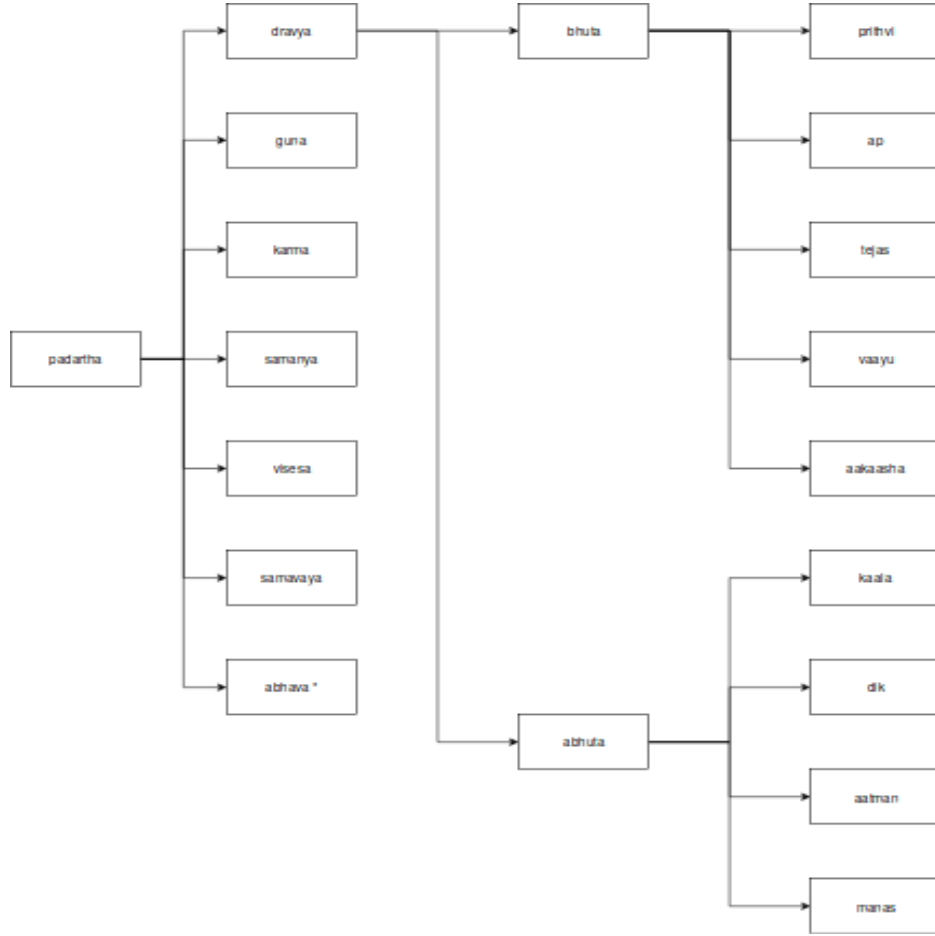


Figure 1: The Vaisheshika ontology

- *karma* (activity): A karma is defined as a transient feature of a substance. Like gunas, karmas do not have a separate existence, and are only manifested by their association to the substances. Karma are only associated with perceivable objects and substances
- *samanya* (generality): Since there are plurality of substances, there will be relations among them. When a property is found common to many substances, it is called samanya.
- *visheha* (particularity): While there is plurality, there is a difference between individual substances, that help identify them differently. When

a property is found different between two similar objects, it is called vishesha.

- *samavaya* (inherence): The relation between the cause and effect is not identified or perceived directly. However, there is a relation between the states of related objects, which is represented by *samavaya*.

The later Vaisheshikas introduced the concept of *abhaava* or absence, which was not represented as any of the above. This was done because each of the above categories can be referenced even in their absence, therefore making it necessary to include it in the ontology. A diagrammatic representation of the Vaisheshika ontology is given in figure 1. For brevity, the types of *guna* are not represented in the figure.

### 3. WordNet

WordNet is an online lexical database, which shows the relations between the entries. These relations are based on the part of speech of the words. One of the most important contributions of WordNet is that it disambiguates the senses of each of the polysemous words in the network. This makes establishing semantic relations between words easier, as each word (with its sense ID) means exactly one thing.

Originally, WordNet has relations which are defined below [4]:

- Synonymy: Synonymy is the relation that establishes that two words mean either exactly or roughly the same. This is based on their meaning in the dictionary. Note that "word" refers only to one particular sense of that word. Synonymy is considered a basic relation, primarily because WordNet bases the concept of synsets as the basic relation of the word senses. A synset of a word is a set of all the words which are synonyms of it. This is primary is sense disambiguation. Synonymy is applied to all the parts of speech in WordNet, which are nouns, adjectives, verbs and adverbs.
- Antonymy: Antonymy is the relation that establishes that two words are opposite. The opposite is based on the synset of the word, such that the synset of a word are antonyms of the synset of its antonym. Antonyms are also valid for all the parts of speech.

**WordNet Search - 3.1**  
- [WordNet home page](#) - [Glossary](#) - [Help](#)

Word to search for:

Display Options:

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations  
Display options for sense: (gloss) "an example sentence"

**Noun**

- [S:](#) (n) **wordnet** (any of the machine-readable lexical databases modeled after the Princeton WordNet)
- [S:](#) (n) **WordNet**, [Princeton WordNet](#) (a machine-readable lexical database organized by meanings; developed at Princeton University)

Figure 2: WordNet UI

- **Hyponymy:** The subordinate relationship, a hyponym of a word denotes the 'type-of' relationship between two words (if A is a hyponym of B, then A is a type of B). This is a relationship restricted to nouns.
- **Hypernymy:** The superordinate relationship is called the hypernymy relationship. The hypernymy relationship is the 'class-of' relationship between two words. Hypernymy is the opposite of the hyponymy relationship.
- **Meronymy:** Meronymy is a complex semantic relation. WordNet distinguishes component, substantive and member parts as the meronym relation. Therefore, a meronym can be considered as a 'part-of' relationship.
- **Holonymy:** Holonymy is the opposite of meronymy, which can be considered the 'whole-of' or 'class-of' relationship. Both meronymy and holonymy are relationships that apply to nouns only.
- **Troponymy:** Troponyms are verb-verb relationships, where the verbs are similar actions to each other, or can be used in a similar context.

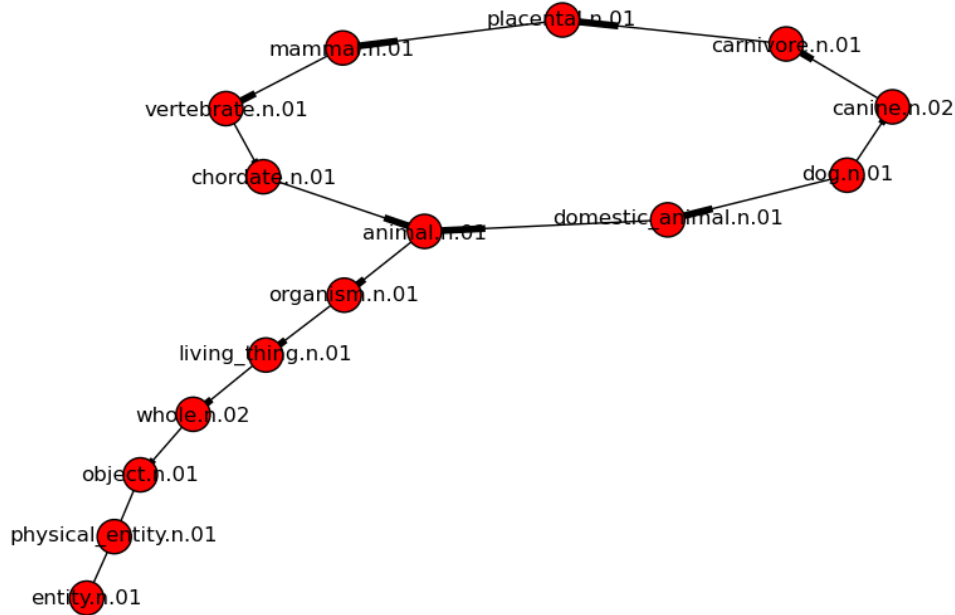


Figure 3: A Graphical Representation of a WordNet Snippet

Troponymy are similar to homonymy, but can be based on the specificity of the use of the verb.

- Entailment: An entailment relationship is a verb-verb relationship that denotes a tautological relationship between one action performed by one participant which directly implies that another action is being performed by another participant of the former action. These are quite common in transitive verb pairs

Figure 3 is a graphical representation of some a snippet of the WordNet structure, the nodes representing the word senses, while the edges represent the relation of synonymy. Note that WordNet is a fully connected hypergraph, which means that it is possible to represent only one relation as a graph. This hypergraph structure has been used in multiple other projects, as WordNet has inspired multiple other knowledge representations including BableNet, a very large semantic network made by combining WordNet and

the Wikipedia articles hypergraph [9], VerbNet, a comprehensive lexicon of verbs and their relations [5], Freebase [10] and DBPedia [11].

#### 4. Ontology for this Project

The ontology used for this project is an adaptation of the Vaisheshika ontology, which is based on the words chosen from the Ashtadhyayi.

##### 4.1. Challenges in Ontology Choice

The semantic conditions chosen for the project were lexicographically chosen from the Ashtadhyayi. Therefore, while I would have wanted to implement a word oriented lexical resource similar to WordNet and added more conceptual information on it, it was not possible to do so, simply because such a graph would be too sparse to represent any useful information. The Vaisheshika ontology, however, suited the words chosen.

It must be noted that the Vaisheshika ontology is very elaborate in its classification of words, as it does so for a large portion of the Sanskrit language. However, this was not the case with the I had to develop. The words chosen to represent the various classifications in the ontology were not chosen from the Ashtadhyayi conditions, they are English classes, as no such semantic constraints could be created otherwise. Therefore a very contracted and modified classification of words is presented in this project.

##### 4.2. Ontology Design

I have designed a simple ontology, as shown in the figure 4. The semantic categorization of words is based on the nature of their existence, and based on their meanings according to that. The meanings of the chosen words were originally chosen from the Monier-Williams Sanskrit to English Dictionary [12], verified from the Ashtadhyayi of Panini [13]. The most basic unit of the ontology is called an 'entity'. Each entity has a form (which can be dependent on the encoding and the type of font) and a meaning as stored in the above references. An example of this basic concept is provided below in figure 5.

The structure of the ontology is as follows (refer figure 4):

- Object: An object is similar to the category of substance in the original Vaisheshika ontology. A member of a class object has an independent existence, which is intransient. An object can be modified by

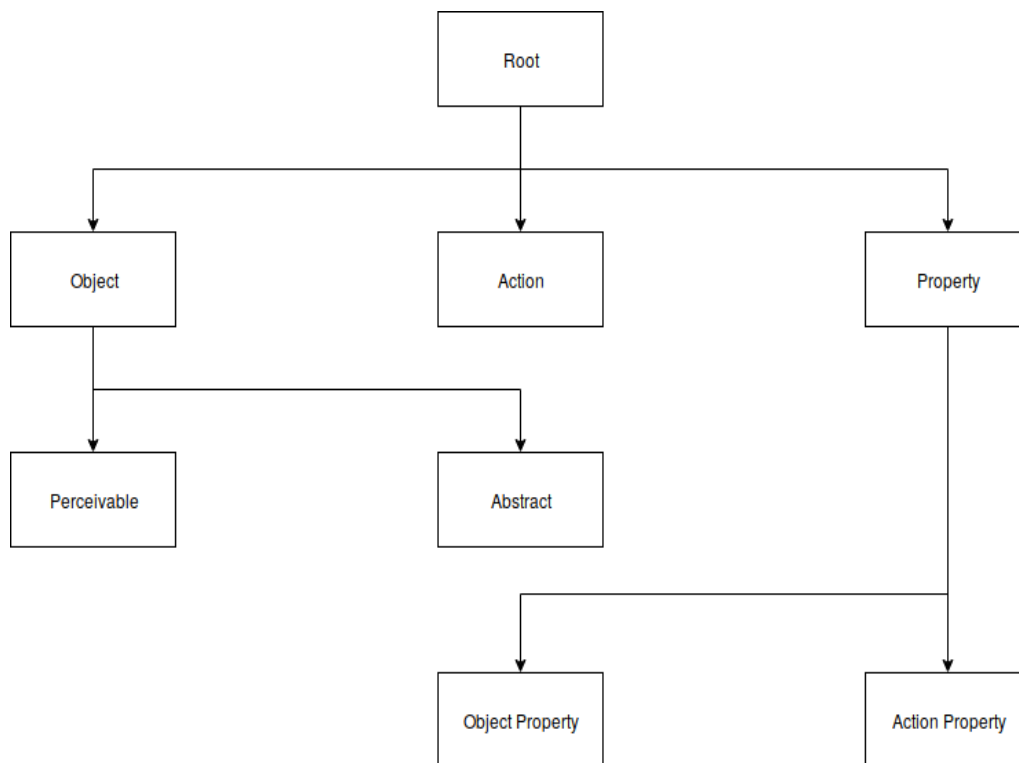


Figure 4: The Structure of my Ontology

```

<entity>
  <form> Silpin </form>
  <meaning> belonging to or skilled in art </meaning>
</entity>

```

Figure 5: Basic XML notation



a property, and there is a link between the property and the object of modifier-modified. This is subclassified in a manner similar to the Vaisheshika definition of *artha* and *budhyapeksham*, except very particularly for objects only.

- Perceivable: A perceivable object is one which can be cognized by one or a combination of the senses. While the nomenclature of a perceivable object (and its properties) are abstractly defined, it serves as a referent to an entity which can be
- Abstract: An abstract object is an intellectually cognizable entity with no perceivable manifestation. Thoughts, ideas, sentience and such concepts are classified as abstract objects.
- Action: An action is a transient property of an object, on the basis of which it undergoes some change in its properties. An object can be said to participate in or perform an action (irrespective of its sentience).
- Property: A property is an intransient modifier of an object or action. There are entities that describe the property of the object which are not inherent to the object itself. I also include in this category entities which describe the nature of an action.

This ontology, while not entirely complete, was enough to categorize all the words of the chosen section of the Ashtadhyayi.

## 5. Conclusion

The semantic conditions I have chosen are very broadly defined, and therefore served to classify most words into one category. There are exceptions, which are as follows:

- A entity can denote an action while being an abstract object. Since the dictionary wasn't based on parts of speech, distinguishing this was difficult. It is the semantic conditions that explain which possible affixes can be applied to the word form, which was used to solve this.
- A property can describe or modify both an action and an object. Since actions and abstract objects can be modified by a property. This was still difficult to determine, and the Ashtadhyayi translations are too difficult to read and understand, compounding the problem.

The ontological structure however, is fairly extensible. The XML structure can be given attributes for further description. However, extending the ontology will make it similar to the Vaisheshika ontology.

## References

- [1] F. Moltmann, Natural language ontology, in: Oxford Research Encyclopedia of Linguistics, 2017.
- [2] C.-R. Huang, Ontology and the lexicon: A natural language processing perspective, Cambridge University Press, 2010.
- [3] J. A. Bateman, The theoretical status of ontologies in natural language processing, arXiv preprint cmp-lg/9704010 (1997).
- [4] G. A. Miller, Wordnet: a lexical database for english, Communications of the ACM 38 (1995) 39–41.
- [5] K. K. Schuler, Verbnet: a broad-coverage, comprehensive verb lexicon (2005).
- [6] H. Liu, P. Singh, Conceptnet: a practical commonsense reasoning tool-kit, BT technology journal 22 (2004) 211–226.
- [7] D. H. H. Ingalls, Materials for the study of navya-nyaya logic (1951) 37–39.
- [8] B. K. Matilal, Nyāya-Vaiśeika, Otto Harrassowitz Verlag, 1977.
- [9] R. Navigli, S. P. Ponzetto, Babelnet: Building a very large multilingual semantic network, in: Proceedings of the 48th annual meeting of the association for computational linguistics, Association for Computational Linguistics, pp. 216–225.
- [10] K. Bollacker, C. Evans, P. Paritosh, T. Sturge, J. Taylor, Freebase: a collaboratively created graph database for structuring human knowledge, in: Proceedings of the 2008 ACM SIGMOD international conference on Management of data, AcM, pp. 1247–1250.
- [11] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, Z. Ives, Dbpedia: A nucleus for a web of open data, in: The semantic web, Springer, 2007, pp. 722–735.

- [12] M. Monier-Williams, Sanskrit-English Dictionary, , 1964.
- [13] Pāṇini, S. C. Vasu, The Ashtadhyāyī, Motilal Banarsidass, 1962.