

– “synonym”, “as known as”, “to be known as” hoặc “see”: thể hiện quan hệ *đồng nghĩa* giữa các khái niệm.

barrel vault, also known as a tunnel vault: an architectural element formed by the extrusion of a single curve along a given distance

– “see also”: thể hiện các khái niệm *có liên quan* với nhau.

anamorphic: An optical system which has different magnifications in the vertical and horizontal dimensions of the picture. **See also aspect ratio, contrast with spherical.**

– “sub class of”: thể hiện mối quan hệ *đặc trưng hóa*. (quan hệ *cha-con* giữa các khái niệm).

eutherians: a **subclass of mammals** which give birth to well-developed young. Humans are part of this subclass.

Ngoài ra, một số mẫu cú pháp ngôn ngữ khác cũng rất hữu ích trong việc xác định các quan hệ giữa các khái niệm. Ví dụ, “*concept₁ such as concepts₂*” ngụ ý rằng *concept₂* là 1 khái niệm con của *concept₁*. Do đó, bằng cách sử dụng các quan hệ đơn giản này giữa các khái niệm, cho phép xác định nhiều quan hệ trong ontology.

3.2 WordNet

WordNet là một cơ sở dữ liệu từ vựng lớn nhất của tiếng Anh, được sử dụng như là một nguồn tài nguyên quan trọng trong rất nhiều ứng dụng về xử lý ngôn ngữ tự nhiên và trong các lĩnh vực khác có liên quan. WordNet cho phép ta truy xuất một cách dễ dàng đến các khái niệm và một tập quan hệ rất phong phú giữa các khái niệm như synonyms,

hyponymy, hypernymy, meronymy,... Do đó, cho một cặp khái niệm, ta có thể kiểm tra xem một mối quan hệ nào đó có tồn tại giữa chúng hay không. Nói cách khác, cho một khái niệm *C*, nếu một khái niệm *C'* nằm trong tập *hypernym* hoặc *meronym* của *C*, hoặc các quan hệ ngữ nghĩa khác thì ta có thể thêm mối quan hệ tương ứng vào ontology.

3.3 Link Grammar

Quan sát cú pháp của một câu, ta có thể dễ dàng thấy được nếu hai khái niệm xuất hiện trong cùng một câu thì động từ liên kết hai khái niệm này thường thể hiện cho quan hệ giữa hai khái niệm. Để thực hiện phân tích câu và tìm động từ thể hiện mối quan hệ giữa các khái niệm, chúng ta cần phải phân tích cú pháp (syntactic) và phụ thuộc (dependency). Trong nghiên cứu này, chúng tôi đề xuất sử dụng Link Grammar (D. Temperley and D. Sleator, 1993), một trong những kỹ thuật sử dụng rộng rãi nhất cho việc phân tích cú pháp câu. Link Grammar không chỉ tạo ra cây cú pháp như POS tagger mà còn cung cấp thông tin về sự phụ thuộc giữa các cặp từ trong câu dưới dạng liên kết (quan hệ). Trong nghiên cứu này, chúng tôi sử dụng PTQL (L. Tari *et al.*, 2010) thay sử dụng trực tiếp Link Grammar. Ngoài việc hỗ trợ xây dựng cây cú pháp và các liên kết, PTQL còn hỗ trợ lưu trữ các thông tin này vào cơ sở dữ liệu quan hệ và ngôn ngữ truy vấn để chúng ta có thể thao tác (sửa đổi, bổ sung, truy vấn) dữ liệu một cách dễ dàng.

3.4 Giải thuật xây dựng ontology tự động từ bảng chú giải

Giải thuật xây dựng ontology tự động từ bảng chú giải được mô tả bằng mã giả (pseudo code) trong Giải thuật 1.