

TÍCH HỢP NGŨ NGHĨA SỬ DỤNG CÁC ONTOLOGY HÌNH THỨC

NGUYỄN KIM ANH

Khoa Công nghệ thông tin, Trường Đại học Bách khoa Hà Nội

Abstract. This paper deals with the problem of handling semantic heterogeneity during schema integration. We present an approach on the basis of formal ontologies presented in a logical language to identification of relationship between different local schemas. Semantic similarity relations between definitions in formal ontologies are defined, verified and used for merging formal ontologies. It is found that the result of the merging process can be used for global schema generation.

Tóm tắt. Bài báo đề cập đến vấn đề giải quyết sự không thuần nhất về ngữ nghĩa trong quá trình tích hợp sơ đồ, và một cách tiếp cận dựa trên các ontology hình thức được biểu diễn bởi một ngôn ngữ logic để nhận diện sự tương ứng giữa các sơ đồ địa phương khác nhau. Các quan hệ tương tự về ngữ nghĩa giữa các định nghĩa trong các ontology hình thức được xác định, được kiểm tra và được sử dụng để trộn các ontology hình thức. Kết quả của quá trình trộn có thể được sử dụng để sinh ra sơ đồ tổng thể.

1. GIỚI THIỆU

Việc thiết kế các hệ thống thông tin lớn thường kéo theo sản sinh ra một số lớn các sơ đồ địa phương khác nhau. Các sơ đồ địa phương này cần được tích hợp để hình thành nên một sơ đồ tổng thể. Một trong những trở ngại lớn đối với nhiệm vụ này là sự không thuần nhất về ngữ nghĩa giữa các sơ đồ địa phương. Các ngữ nghĩa được nói tới ở đây ám chỉ các mối quan hệ giữa các thuật ngữ được sử dụng để đặt tên cho các phần tử trong sơ đồ và các dữ liệu mà chúng biểu diễn hay sự thông dịch của các thuật ngữ này. Trong [1, 3, 7], một số tác giả đã đề cập đến khả năng sử dụng một số tri thức về ngữ nghĩa của các sơ đồ cơ sở dữ liệu (CSDL), biểu diễn chúng thông qua một logic mô tả để so sánh các khái niệm. Tuy nhiên, các cách tiếp cận này không đề cập đến cách giải quyết trường hợp đặt tên không thống nhất trong các sơ đồ CSDL địa phương. Chúng tôi cho rằng, các ontology hình thức phải chứa nhiều tri thức hơn và phải bao hàm cả một số tri thức về ngôn ngữ liên quan đến các thuật ngữ được giải nghĩa trong ontology. Các ontology hình thức cho phép định nghĩa một cách hình thức và tường minh các tri thức và ngữ nghĩa của các thuật ngữ có thể là một giải pháp tiềm tàng cho sự không thuần nhất về ngữ nghĩa trong quá trình tích hợp sơ đồ. Một ontology hình thức bao gồm một tập các tiền đề logic nhằm cung cấp ý nghĩa của các thuật ngữ đối với một cộng đồng. Một tập các tiền đề logic đối với một thuật ngữ là một định nghĩa tiềm ẩn của thuật ngữ và chỉ có một định nghĩa tiềm ẩn cho một thuật ngữ đối với mỗi cộng đồng. Các định nghĩa tiềm ẩn này phải mô tả được các đặc trưng của miền ứng dụng và những giả thiết tiềm ẩn trong việc thiết kế CSDL, trong cách đặt tên các đối tượng đối với miền ứng dụng đó. Các ontology hình thức được xây dựng dựa trên sự hiểu

biết của các thành viên trong một cộng đồng và giúp cho việc giảm thiểu sự nhập nhằng khi trao đổi với các cộng đồng khác. Chắc chắn rằng, các ontology này mang nhiều tri thức hơn so với các định nghĩa sơ đồ trong các CSDL vì các sơ đồ chỉ liên quan chính đến việc tổ chức dữ liệu trong các CSDL phù hợp với sự đặc tả các yêu cầu ứng dụng. Tuy nhiên, các định nghĩa sơ đồ không độc lập đối với các định nghĩa ontology và ngược lại, chúng bao hàm một phần tri thức của ontology đối với một cộng đồng.

Trong bài báo này, chúng tôi sẽ chỉ ra các ontology hình thức có thể được sử dụng để nhận diện sự tương ứng giữa các sơ đồ địa phương khác nhau. Các ontology này được trộn lại dựa trên các quan hệ tương tự sau khi đã được kiểm tra là nhất quán với các ngữ nghĩa của các sơ đồ địa phương. Ontology trộn kết quả có thể được sử dụng để xây dựng sơ đồ tổng thể cho một hệ CSDL liên hiệp ghép nối chặt. Nội dung của bài báo được trình bày như sau: Mục 2 đưa ra một định nghĩa cho ontology hình thức và ngôn ngữ được lựa chọn để biểu diễn nó, Mục 3 trình bày một phép dịch ngữ nghĩa các sơ đồ địa phương cho phép xây dựng các ontology địa phương, Mục 4 trình bày các bước cho phép trộn các ontology địa phương để sinh ra ontology tổng thể. Cuối cùng, Mục 5 đưa ra một số nhận xét và kết luận.

2. CÁC ONTOLOGY HÌNH THỨC VÀ NGÔN NGỮ BIỂU DIỄN

Định nghĩa 1. Một ontology hình thức được định nghĩa là bộ bảy $O := (C, ROOT, H, R, L, F, G)$ trong đó:

C là một tập các khái niệm. Mỗi một khái niệm $C \in C$, không phải là các kiểu dữ liệu nguyên tố, có một mô tả hay một định nghĩa trong ontology này và được đặt một cách tương ứng vào sự phân loại H của ontology.

$ROOT$ là một khái niệm đỉnh đặc biệt, $ROOT$ không thuộc C nhưng trong phân loại H , nó là khái niệm trên mọi khái niệm khác trong C .

H là một sự phân loại. Các khái niệm được phân loại bởi quan hệ phân loại H có tính chất bắc cầu và phi chu trình. $H \subset C \times (C \cup \{ROOT\})$, trong đó $H(C, C')$ có nghĩa là C là một khái niệm con của C' và $\forall C \in C : H(C, ROOT)$ là đúng.

R bao gồm một tập các quan hệ R^A để mô tả các đặc tính của một khái niệm và một tập các vai trò ngữ nghĩa R^T để xác định vai trò ngữ nghĩa của một khái niệm đối với khái niệm đang được mô tả. Nói chung, một quan hệ được tham chiếu đến bởi một số từ vựng của nó và nó đặc tả một cặp (C, C') với $C, C' \in C$. Một quan hệ R đặc tả một cặp (C, C') bổ sung thêm vào sự mô tả đối với khái niệm C một quan hệ R có hạn chế miền giá trị là khái niệm C' , có nghĩa là, một thể hiện c của C có thể có quan hệ R đến một thể hiện khác c' chỉ nếu $c' \in C'$. Các hàm d và r được áp dụng đối với R cho biết các khái niệm miền xác định và miền giá trị C và C' tương ứng của R . Tập vai trò ngữ nghĩa đối với một ontology phụ thuộc vào miền ứng dụng cụ thể đang được xem xét và các tri thức về ngôn ngữ của các thuật ngữ được sử dụng trong miền ứng dụng này.

L là một bộ từ vựng bao gồm một tập các từ vựng đối với các khái niệm, kí hiệu là L^C và một tập các từ vựng đối với các quan hệ, kí hiệu là L^{RA} , có nghĩa là $L := L^C \cup L^{RA}$.

F, G là các hàm tham chiếu với $F : C \rightarrow 2^L$ và $G : R^A \rightarrow 2^L$. F và G móc nối mỗi khái niệm hay quan hệ đến một tập các từ vựng tương ứng trong ontology đã cho. Nói chung, một khái niệm hay một quan hệ có thể được tham chiếu đến bởi một số từ vựng có cùng nghĩa hay có liên quan một cách gần gũi. Chúng ta giả thiết rằng $\forall C \in C$ thì $F(C)$ chứa C

và $\forall R \in \mathbf{R}$ thì $\mathbf{G}(R)$ chứa R .

Sau đây, chúng ta sẽ thảo luận về ngôn ngữ được sử dụng để biểu diễn các khái niệm trong một ontology hình thức. Chúng tôi đưa ra các tính chất cần có đối với một ngôn ngữ được lựa chọn để mô tả các khái niệm:

- Ngôn ngữ này phải là một ngôn ngữ khai báo vì các ontology điển hình được sử dụng để biểu thị các quan hệ ràng buộc đối với các khái niệm. Bên cạnh đó, bản chất khai báo của ngôn ngữ sẽ tạo điều kiện dễ dàng thực hiện các suy diễn dựa trên các định nghĩa của nó.
- Ngôn ngữ này phải có khả năng biểu thị các khái niệm thông qua một tập các vai trò, mỗi vai trò mô tả một khía cạnh đặc biệt của thông tin cần được biểu diễn trong CSDL.
- Ngôn ngữ này phải có các phép toán trong thế giới thực được mô hình hóa (như xác định khái niệm con của hai khái niệm, đối sánh mẫu,...) mà có thể có ích trong việc đối sánh và thao tác trên các định nghĩa của khái niệm.

Với những tính chất cần thiết đối với ngôn ngữ biểu diễn các khái niệm trong một ontology, chúng tôi sử dụng cùng một logic mô tả, được gọi là ALCQI, để biểu diễn các khái niệm trong ontology của các sơ đồ địa phương và ontology trộn. ALCQI là một ngôn ngữ biểu diễn khá mạnh bao gồm tất cả các cấu trúc khái niệm thông thường như hợp, giao, phủ định, các hạn chế số với lượng từ \exists, \forall và quan hệ ngược.

Sau đây, chúng tôi đưa ra một phép dịch ngữ nghĩa các sơ đồ địa phương cho phép xây dựng các ontology địa phương.

3. XÂY DỰNG ONTOLOGY HÌNH THỨC CHO CÁC SƠ ĐỒ ĐỊA PHƯƠNG

Để xây dựng các ontology hình thức cho các sơ đồ địa phương, chúng ta có thể sử dụng các sơ đồ thực thể-liên kết được mở rộng với các vai trò ngữ nghĩa hay các đồ thị khái niệm mở rộng như trong [5], ở đây, các liên kết thực thể cũng được xem như một khái niệm và các vai trò ngữ nghĩa của các thực thể kéo theo liên kết sẽ được sử dụng để định nghĩa cho khái niệm biểu diễn liên kết đó.

Trong phần này, chúng tôi sẽ chỉ ra rằng, các ngữ nghĩa được phản ánh trong sơ đồ thực thể-liên kết đối với các CSDL địa phương có thể được nắm bắt trong ontology địa phương thông qua một phép dịch từ sơ đồ thực thể-liên kết thành các định nghĩa tiềm ẩn hay các tiền đề thuật ngữ trong logic mô tả ALCQI.

Chúng tôi kết hợp một khái niệm nguyên tố A đối với mỗi miền giá trị thuộc tính hay mỗi thực thể A , một vai trò nguyên tố P đối với mỗi thuộc tính P và mỗi vai trò ngữ nghĩa trong một liên kết R n -ngôi (kéo theo n thực thể) có m thuộc tính liên kết. Đối với logic mô tả ALCQI, chúng tôi kí hiệu T là khái niệm đỉnh, \perp là khái niệm đáy, Π là phép giao và \cup là phép hợp. Các tiền đề thuật ngữ trong logic mô tả ALCQI được suy ra từ một sơ đồ thực thể-liên kết S được xác định như sau:

+ Với mỗi cặp hai thực thể E và F , nếu E là một F trong S , chúng ta có: $E \subseteq F$ với E và F là các khái niệm nguyên tố ứng với các thực thể E và F . Ngược lại, chúng ta có $E \Pi F = \perp$.

+ Với mỗi thực thể E có các thuộc tính A_1, A_2, \dots, A_k với các miền D_1, D_2, \dots, D_k tương ứng, ta có $E = \exists A_1.D_1 \Pi \dots \Pi \exists A_k.D_k$.

+ Với mỗi liên kết n -ngôi R giữa n thực thể E_1, \dots, E_n có m thuộc tính liên kết T_1, \dots, T_m với các miền D_1, D_2, \dots, D_m , tương ứng, ta có:

$$R = \exists T_1.D_1 \Pi \dots \Pi \exists T_m.D_m \Pi \exists R_1.E_1 \Pi \dots \Pi \exists R_n.E_n,$$

ở đây, R_i là vai trò ngữ nghĩa của E_i trong liên kết R .

Để đối sánh các thuật ngữ được sử dụng không thống nhất trong các sơ đồ địa phương, các ontology hình thức này cần giải nghĩa chi tiết hơn về ý nghĩa của các thuật ngữ được sử dụng trong ngữ cảnh cụ thể của các nguồn dữ liệu địa phương.

4. TRỘN CÁC ONTOLOGY

Một bước quan trọng trong quá trình tích hợp các CSDL không thuần nhất là xác định các mối quan hệ giữa các phần tử sơ đồ từ các CSDL địa phương khác nhau và một tích hợp có hiệu lực phải dựa trên một sự hiểu biết đúng đắn ngữ nghĩa của các phần tử sơ đồ. Điều này có nghĩa là, nó phải xác định được liệu các phần tử từ các sơ đồ khác nhau có tham chiếu đến cùng các thực thể trong thế giới thực hay chúng là khác nhau và nếu có thể thì chúng khác nhau ở mức nào?

Cách tiếp cận của chúng tôi dựa trên các ontology địa phương đối với các CSDL địa phương và tích hợp sơ đồ sẽ dựa trên phép trộn các ontology này. Do vậy, nhiệm vụ chính trong việc trộn các ontology là đối sánh các định nghĩa tiềm ẩn từ các ontology hình thức khác nhau. Việc đối sánh được thực hiện dựa trên các quan hệ tương tự, cụ thể, chúng tôi xét các quan hệ sau: \equiv (bằng, tương đương ngữ nghĩa), \leq (đặc biệt hóa), \sim (bao trùm).

Cho hai ontology $O = (C, ROOT, H, R, L, F, G)$ và $O' = (C', ROOT', H', R', L', F', G')$. Hệ sẽ dò tìm và phát hiện các quan hệ tương tự giữa hai khái niệm và các quan hệ tương tự giữa hai quan hệ thuộc các ontology khác nhau.

Định nghĩa 2. Cho hai khái niệm $C \in C$, $C' \in C'$ và hai quan hệ $R \in R$, $R' \in R'$. Ta định nghĩa:

$C \equiv C'$ nếu và chỉ nếu $F(C) \cap F'(C') \neq \emptyset$ và $C \equiv C'$.

$C \leq C'$ nếu và chỉ nếu $F(C) \cap F'(C') \neq \emptyset$ và $C \Pi C' \equiv C$.

$C \sim C'$ nếu và chỉ nếu $C \Pi C' \neq \perp$.

$R \equiv R'$ nếu và chỉ nếu $G(R) \cap G'(R') \neq \emptyset$ và $r(R) \equiv r(R')$.

$R \leq R'$ nếu và chỉ nếu $G(R) \equiv G'(R') \neq \emptyset$ và $r(R) \leq r(R')$.

Các trường hợp còn lại là rời nhau hay không có quan hệ tương tự với nhau.

Định nghĩa 3. Cho hai ontology O và O' . Ta nói O là không mâu thuẫn về ngữ nghĩa đối với O' nếu không tồn tại hai khái niệm C thuộc O và C' thuộc O' sao cho $F(C) \cap F'(C') \neq \emptyset$ và $C \Pi C' = \perp$.

Ví dụ 1. Ví dụ sau đây chỉ ra một phần của các định nghĩa trong các ontology được sử dụng trong quá trình trộn ở phần sau.

Trong ontology O , ta có các định nghĩa tiềm ẩn của các khái niệm như sau:

Đạy = \exists Agnt.GiảngViên $\Pi \exists$ Obj.MônHọc $\Pi \exists$ Rcpt.SinhViên.

GiảngViên = Người $\Pi \exists$ HọcVi. String $\Pi (\neg$ SinhViên).

Người = \exists Tên. String $\Pi \exists$ ĐịaChỉ. KiểuĐC.

KiểuĐC = \exists Tên. TP. String.

MônHọc = \exists Tên. String $\Pi \exists$ SốTiết. Int $\Pi (\neg$ GiảngViên) $\Pi (\neg$ SinhViên).

SinhViên = Người Π NgànhHọc.String Π (\neg GiảngViên).

...

Trong ontology O' , ta có các định nghĩa tiềm ẩn của các khái niệm như sau:

ThuyếtTrình = \exists Agnt.GiáoSư Π Obj.ChuyênĐề Π Rcpt.SVTN.

GiáoSư = \exists HọTên.String Π ChỗỞ.KiểuChƠ Π HọcVị.String Π (\neg SVTN).

KiểuChƠ = \exists SốNhà.Int Π (TênPhố.String Π TênTP.String).

ChuyênĐề = \exists Tên.String Π Loại.Int Π (\neg GiáoSư) Π (\neg SVTN).

SVTN = \exists Tên.String Π (ChỗỞ.KiểuChƠ Π (\exists NgànhHọc.String Π (\exists ĐềTàiTN. String Π (\neg GiáoSư)

...

Ở đây, ta kí hiệu các vai trò ngữ nghĩa như sau: Agnt là tác nhân, Obj là đối tượng chịu tác động và Rcpt là đối tượng nhận tác động.

Hệ lập luận xác định được các quan hệ tương tự sau:

Tên \equiv HọTên, KiểuChƠ \leq KiểuĐC, ChỗỞ \leq ĐịaChỉ, GiáoSư \leq GiảngViên, SVTN \leq SinhViên, ChuyênĐề \sim MônHọc, ThuyếtTrình \sim Day...

Hai ontology O và O' không mâu thuẫn về ngữ nghĩa có thể được trộn để xây dựng ontology tổng thể.

Thuật toán 1. Trộn hai ontology địa phương O và O' để xây dựng ontology tổng thể Og .

Vào: $O = (C, ROOT, H, R, L, F, G)$ và $O' = (C', ROOT', H', R', L', F', G')$ không mâu thuẫn về ngữ nghĩa.

Ra: $Og = (Cg, ROOTg, Hg, Rg, Lg, Fg, Gg)$.

Cách làm:

+ Đặt $Lg = L \cup L'$, $Rg = R \cup R'$ và $Gg = G \cup G'$. Cg, Hg và Fg ban đầu được đặt là rỗng.

+ Duyệt các phân loại H và H' trong các ontology địa phương từ trên-xuống bắt đầu từ các khái niệm trên đỉnh của chúng. Với mỗi khái niệm C trong $O(O')$, ta thêm một khái niệm mới C vào Cg và $Fg(C) = F(C)(F'(C))$. Quan hệ phân loại của khái niệm mới với một siêu khái niệm đã tồn tại trong ontology tổng thể được thiết lập khi nó được thêm vào.

+ Khi thêm một khái niệm mới C' , nếu trong ontology tổng thể đã có một khái niệm C mà $C \equiv C'$, ở đây giả sử $C \in O$ và $C' \in O'$, khái niệm mới này sẽ không được thêm vào mà chỉ lưu giữ lại thông tin về nó bằng cách thay thế C bởi khái niệm $CC' = C \sqcup C'$ và đặt $Fg(CC') = F(C) \cup F'(C')$. Như vậy, để thêm một khái niệm mới C' vào ontology tổng thể, điều kiện sau đây phải được thỏa mãn $\forall c \in Og$ thì $\neg[C' \equiv c]$.

+ Quan hệ đặc biệt hóa (\leq) của các khái niệm thuộc các ontology khác nhau được thêm vào như một quan hệ phân loại trong Hg . Ta thiết lập một khái niệm con với mỗi khái niệm tổng quát hóa của nó trong Hg với điều kiện khái niệm con này không thể suy diễn ra được từ sự phân loại trong ontology tổng thể.

+ Nếu hai khái niệm $C \in O$, $C' \in O'$ mà $C \sim C'$, một khái niệm mới $D = C \Pi C'$ có thể được thêm vào ontology tổng thể với $Fg(D) = F(C) \cap F'(C')$. Tuy nhiên, những trường hợp như thế này xảy ra khá thường xuyên trong quá trình trộn các ontology. Do vậy, cần phải có sự giám sát của con người vào thời điểm xem xét những trường hợp này và có thể quyết định về sự cần thiết tạo các khái niệm giao như vậy.

+ Nếu hai khái niệm là bao trùm hay rời nhau mà chúng có một siêu khái niệm chung thì một khái niệm dựa trên siêu khái niệm chung này có thể được định nghĩa trong ontology tổng

thể.

Mệnh đề 1. *Ontology tổng thể Og chứa hai ontology địa phương O và O' và tất cả các quan hệ ngữ nghĩa trong O và O' đều được bảo toàn trong Og .*

Chứng minh: Với đầu vào là hai ontology O và O' không mâu thuẫn về ngữ nghĩa, Mệnh đề 1 sẽ được đảm bảo vì dựa trên các bước tiến hành của thuật toán trộn hai ontology địa phương, với mỗi khái niệm trong ontology địa phương, ta đều tạo một khái niệm tương ứng trong ontology tổng thể. Điều này đảm bảo rằng mỗi khái niệm trong ontology địa phương đều đã được biểu diễn bởi một khái niệm trong ontology tổng thể mà điều này là quan trọng đối với việc xây dựng sơ đồ tổng thể. Các quan hệ phân loại thuộc các ontology khác nhau cũng được thiết lập như một quan hệ phân loại trong Hg . Cuối cùng, Fg cũng chứa cả F và F' . ■

Ví dụ 2. Hai ontology O và O' trong Ví dụ 1 là không mâu thuẫn về ngữ nghĩa. Thực hiện thuật toán trộn ontology, ta có Og với Cg và Hg như sau:

$Cg = \{\text{Người, GiảngViên, SinhViên, GiáoSư, SVTN, TênHọTên, KiểuChƠ, KiểuĐC, ChỗỞ, ĐịaChỉ, ChuyênĐề, MônHọc, ThuyếtTrình, Dạy, ChuyênĐề, MônHọc, ChuyênĐề(MônHọc,...)}\}.$

$Hg = \{\text{KiểuChƠ} \leq \text{KiểuĐC}, \text{ChỗỞ} \leq \text{ĐịaChỉ}, \text{GiảngViên} \leq \text{Người}, \text{SinhViên} \leq \text{Người}, \text{GiáoSư} \leq \text{GiảngViên}, \text{SVTN} \leq \text{SinhViên}, \text{ChuyênĐề} \leq \text{ChuyênĐỀIIMônHọc}, \text{MônHọc} \leq \text{ChuyênĐỀIIMônHọc}, \dots\}.$

5. KẾT LUẬN

Gần đây, các nhà nghiên cứu và thực hành trong lĩnh vực cơ sở dữ liệu và tích hợp thông tin đã quan tâm đến các cách tiếp cận tích hợp ngữ nghĩa dựa trên ontology [9]. Tuy nhiên, họ thường sử dụng một ontology dùng chung hay sử dụng các giải thuật heuristics, phương pháp học máy để pháp hiện sự tương ứng giữa các nguồn dữ liệu địa phương. Một trong những hạn chế của các tiếp cận này là khó xác định được sự tương ứng về thuật ngữ giữa các nguồn dữ liệu khác nhau [4, 6]. Trong bài này, chúng tôi đã trình bày một cách tiếp cận sử dụng các ontology hình thức như là cơ sở đối với tích hợp và giải quyết các vấn đề không thuần nhất về thuật ngữ, ngữ nghĩa trong quá trình tích hợp các sơ đồ địa phương. Ta sử dụng các định nghĩa tiềm ẩn trong các ontology hình thức để xác định các quan hệ tương tự giữa các khái niệm và các quan hệ. Chất lượng của các ontology hình thức và đặc biệt độ chi tiết của các mô tả trong ontology hình thức đóng một vai trò quan trọng trong cách tiếp cận này. Những quan hệ tương tự xác định được giữa các thuộc tính và các lớp của các sơ đồ địa phương khác nhau có thể được xem như các tri thức giữa các sơ đồ địa phương và có thể được sử dụng đối với việc xử lý và tối ưu hóa các câu hỏi tổng thể. Hy vọng cách tiếp cận này có thể được áp dụng trong quá trình tích hợp các nguồn dữ liệu khác nhau để xây dựng các kho dữ liệu.

TÀI LIỆU THAM KHẢO

- [1] D. Calvanese, G. D. Giacomo, M. Lenzerini, D. Nardi, R. Rosati, Description logic framework for information integration, *Proc. Of Sixth International Conference on PKRR*, Italy, 1998.

- [2] D. Calvanese, M. Lenzerini, D. Nardi, *Logics for Databases and Information Systems*, Kluwer, 1998.
- [3] T. Catarci and M. Lenzerini, Representing and using interschema knowledge in cooperative information systems, *Journal of ICIS* **4** (2) (1993) 375–398.
- [4] AnHai Doan, AY. Halevy, Sematic integration research in the database community, *AI Magazine* **26** (1) (2005) 83–91.
- [5] J. Farques, M. C. Landau, A. Dugourd, and L. Catach, Conceptual graphs for semantics and knowledge processing IBM, *J. Res. Develop* **30** (1) (1986).
- [6] HongHai Do, Ph.D. Thesis: “Schema matching and mapping-based Data integration”, University of Leipzig, 2005.
- [7] V. Kashyap and A. Sheth, Semantic and schematic similarities between database objects: A context-based approach, *International Journal on VLDB* **5** (4) (1996) 276–304.
- [8] B. Moulin and P.N. Creasy, Extending the conceptual graph approach for data conceptual modelling, *Data & Knowledge Engineering* **8** (4) (1992) 223–248.
- [9] N.F. Noy, Semantic integration: A survey of ontology-based approaches, *ACM SIGMOD* **33** (4) (2005) 65–70.
- [10] P. Weinstein, W. Birmingham, Comparing concepts in differentiated ontologies, *Proc. Of KAW*, Canada, 1999.

Nhận bài ngày 9 - 1 - 2006

Nhận lại sau sửa ngày 4 - 9 - 2007