

래티스(lattice) 구조를 이용한 애니메이션 검색 시스템 설계

Design of Retrieval System in Animation Products using Lattice Structure

우선미, 홍정표
전북대학교 산업디자인학과

Woo Seon-Mi, Hong Jung-Pyo
Dept. of Industrial Design, CBNU

• Key words: lattice structure, animation, retrieval, query expansion

1. 서론

정보통신과 인터넷의 발전으로 현대인은 수많은 정보 속에 살고 있으며 여러 분야에서 검색 서비스를 이용하고 있다. 국내 애니메이션 분야에서는 웹(WWW)에서 전 세계의 애니메이션을 검색할 수 있는 서비스가 이루어지지 않고 있으며, 몇몇 상업용 사이트에서만 애니메이션 정보 검색 서비스를 실시하고 있다. 이러한 상업용 사이트에서는 각 사이트마다 취급하는 애니메이션만을 대상으로 서비스를 하고 있고, 분류 또한 취급하는 애니메이션에 따라 그 체계가 다양하다. 그리고 기존 사이트에서 서비스하고 있는 검색 방법들은 키워드 정합매칭(exact matching) 기법을 이용하고 있으므로 사용자가 원하는 검색 결과를 얻기 위해서는 애니메이션 분류나 제작사와 같은 정보를 알고 있어서 키워드를 직접 타이핑해야하고, 타이핑하는 용어 또한 정확해야만 정확한 검색결과를 얻을 수 있다. 분류에 의한 검색의 경우엔, 한가지 분류 기준으로 검색한 결과에서 여러 정보를 한번에 볼 수 없기 때문에 검색 기준을 바꿔가며 여러 번에 걸쳐서 검색해야한다는 단점이 있다. 본 논문에서는 정확한 사전 지식 없이도 원하는 결과를 얻을 수 있고, 여러 분류 기준들을 한번에 적용하여 검색할 수 있도록 하기 위하여 래티스 구조를 적용한 애니메이션 검색 시스템(ARS-lattice : Animation Retrieval System using Lattice)을 제안한다. 논문에서 제안하는 방법을 이용하면 검색 영역의 내용 구조를 미리 알고 있지 않거나 자신의 요구를 용어들의 조합으로써 표현할 수 없는 일반 사용자들이 상세한 질의를 입력하지 않고 브라우저를 자유롭게 항해하면서 원하는 결과를 얻을 수 있다. 그리고 애니메이션의 여러 분류 기준에 해당하는 용어들로 래티스를 구성하므로 사용자가 원하는 만큼 분류 기준을 AND 조합하여 검색할 수 있다.

2. 이론적 고찰

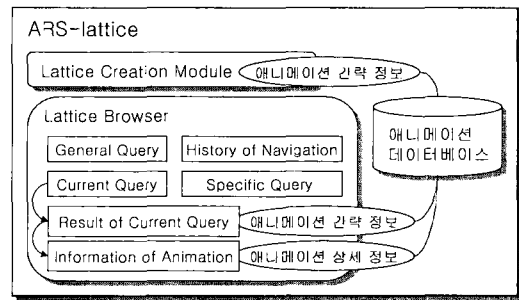
그래프 구조에 기반한 브라우저 환경을 이용한 정보검색에서는 용어를 정점(vertices)로 표현하고 용어들간의 의미론적 관계를 에지(edge)로 표현한다[Salton,1968]. 래티스는 부분순서집합으로 볼 수 있으므로 부분순서집합, $\langle L, \leq \rangle$ 에서, L 에 속하는 모든 원소의 쌍 $\{a, b\} \subseteq L$ 이 최대 하계(the greatest lower bound)와 최소 상계(the least upper bound)를 가진다. 래티스를 이용한 정보검색 방법을 살펴보면, 문서에 부여된 색인어들의 곱(conjunction)을 그래프 내의 각 정점으로 표현해 놓고, 용어 그래프와 문서 부분집합을 브라우징(browsing) 해줌으로써 사용자 질의(query)의 점진적인 확장(gradually enlargement)과 정제(refinement)를 가능하도록 하고 있다[Robert Godin, 1989]. 래티스가 구성되면 래티스의 구조를 브라우징 해주는 기능을 가진 그래픽 사용자 인터페이스를 이용하여 사용자는 원하는 정보를 얻기 위한 방향을 항해(navigation)할 수 있다. 이러한 시스템은 자세한 질의를 주기 위한 지식이 없는 사용자에게 검색 편의성을 제공하는 장점을 지닌다.

3. ARS-lattice의 구성과 래티스 생성

본 논문에서 제안하는 ARS-lattice (Animation Retrieval System using lattice)의 구성은 [그림 3-1]과 같다. 애니메이션 데이터베이스에 입력되어 검색결과 제시될 '애니메이션 상세 정보'는 제목, 장르, 감독, 제작년도, 제작사, 종류, 상영시간, 제작국, 주요 캐릭터 이미지나 주요 이미지이고, 래티스를 구성하기 위해 사용하는 '애니메이션 간략

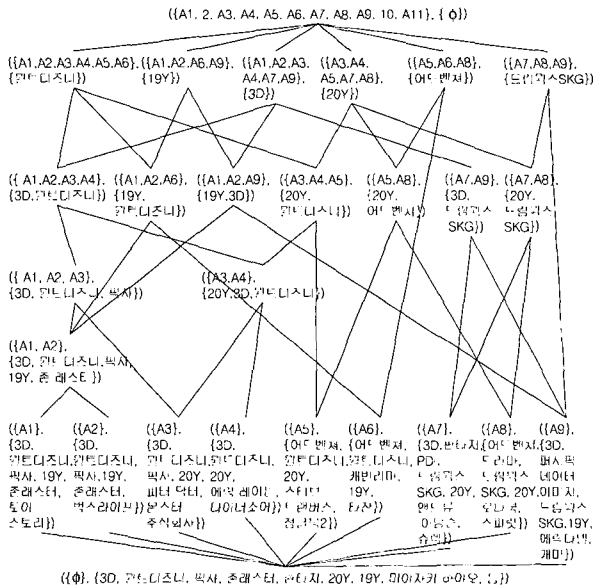
정보'는 제목, 제작사, 감독, 제작 년도, 장르이다. 래티스를 구성하기 위하여 장르는 액션, 드라마, 스포츠, 3D 애니메이션, SF, 환타지, 어드벤처, 다큐멘터리, 뮤지컬로 결정한다.

[그림 3-1] ARS-lattice의 구성



장르는 한 개 이상의 값을 가질 수 있고, 하세다이어그램으로 구성할 때 각 값을 집합에 포함된 하나의 원소로 취급한다. 제작 년도는 1920년대부터 십 년 단위로 12Y(1920년대), 13Y(1920년대)~, 19Y (1990년대), 20Y(2000년대)와 같이 범위를 정한다. 검색결과로 보여지는 정보에서는 년도 범위가 아닌 실제 제작 년도를 표시한다. 'Specific Query'는 하세도표에서 현재 질의가 위치한 정점의 자손 정점들 중에서 바로 아래 단계의 정점들의 개수(질의어 개수)와 각 정점이 나타내는 용어들(질의들), 그리고 각 질의별로 검색된 애니메이션의 개수를 보여준다. 따라서 항상 하세도표의 아래쪽으로부터 보여줌으로써 항해가 계속될수록 검색결과 애니메이션의 개수는 작아진다. 'Current Query'는 'Specific Query'나 'General Query' 또는 'History of Navigation'에서 사용자가 선택한 질의, 질의 개수, 애니메이션의 개수를 보여준다. 'General Query'는 하세도표에서 현재 질의가 위치한 정점의 바로 위 단계의 질의어, 질의 개수, 검색된 애니메이션의 개수를 보여준다. 'History of Navigation'은 사용자가 방문했던 정점들의 리스트를 역순으로 보여줌으로써 원하는 단계부터 다시 검색을 가능하게 해준다. 'Results of Current Query'는 현재 질의에 해당되는 애니메이션들의 간략 정보를 보여준다. 'Information of Animation'에서는 Results of Current Query에서 선택한 애니메이션의 상세 정보를 보여준다. 예를 보이기 위하여 애니메이션 9개로 구성된 하세다이어그램은 [그림 3-1]과 같다. 부분 순서 집합은 방향성 그래프로 나타낼 수 있으나, 부분 순서 집합도형인 하세다이어그램(Hasse diagram)으로 표현하는 것이 편리하여 본 논문에서는 하세다이어그램을 이용하여 래티스를 구성한다. 이 도형은 집합 원소의 개수만큼의 정점을 갖고, $a_i, a_j \in A$ 에 관하여 $a_i \neq a_j$ 이고, $a_i \leq a_j$ 이면 도형에서 a_i 를 a_j 의 아래쪽에 그린다. $a_i \neq a_j$ 이고, $a_i \leq a_j$ 일 때 $a_i \leq a_k \leq a_j$ 이고, $a_i \neq a_k, a_j \neq a_k$ 인 a_k 가 A내에 존재하지 않을 때 a_i 에서 a_j 로 선을 그린다. 부분 순서 관계의 역순서(inverse order)도 부분 순서 관계이므로, 하세다이어그램 상에서 상하로 항해하면서 정점인 질의확장이 가능하다. 래티스의 각 정점들은 애니메이션의 집합 $A \subseteq \mathcal{L}$, 용어들의 집합 $T \subseteq \mathcal{T}$ 로 구성된 완전 쌍(Complete Couple)이다. 이때 \mathcal{L} 는 데이터베이스에 있는 모든 애니메이션을 표현하는 문서의 집합이고, \mathcal{T} 는 래티스를 구성하고있는 모든 용어들의 집합이다.

[그림 3-1] 하세다이어그램

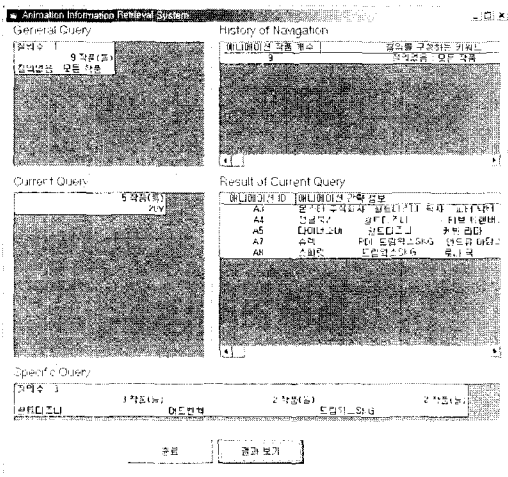


$A_1 < A_2$ 이면 에지 (A_1, A_2) 가 존재하고, 래티스 내에 또 다른 요소 A_3 가 $A_1 < A_3 < A_2$ 와 같이 존재하지 못한다. [그림 3-1]과 같은 하세다이어그램은 자동으로 생성 가능하며, 보통 위에서 아래로 그려나간다.

4. 시나리오

그림 3-1의 래티스를 예로 검색 시나리오를 설명한다. 검색 초기엔 'General Query'에 정보가 없고 'Current Query'엔 모든 문서가 해당된다고 표시되며, 'Specific Query'에는 "질의 수 : 6", 각 정점의 용어(질의어)들인 "윌트 디즈니", "19Y", "3D", "20Y", "어드벤처", "드림웍스 SKG", 그리고 각 질의를 포함하는 작품들의 개수가 표시된다. 'Results of Current Query'에는 'Current Query'에 해당하는 A1~A9의 간략 정보가 표시된다. 이 상태에서 'Specific Query'의 "20Y"를 선택하면 [그림 4-1]과 같이 'Current Query'에 애니메이션 개수인 "5작품(들)"과 질의어인 "20Y"가 표시된다.

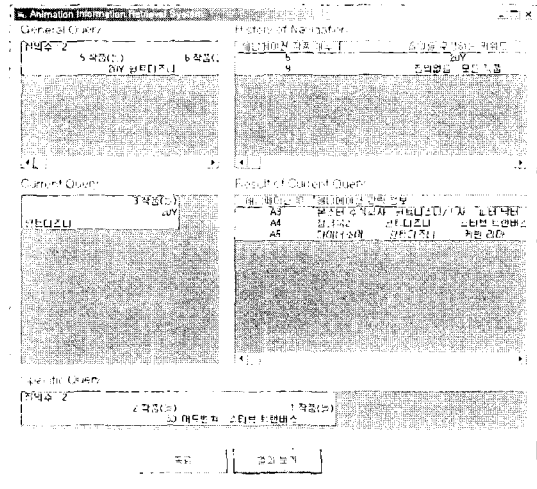
[그림 4-1] 현재 질의가 "20Y"인 경우



'Specific Query'에는 "20Y"의 하위 정점들의 개수인 "질의 수 : 3"과 현 질의어를 제외한 하위 질의어들인 "윌트디즈니", "어드벤처", "드림웍스SKG"가 표시되고, 각 질의어에 해당하는 작품들이 개수가 표시된다. 'Results of Current Query'에는 현재 질의어에 해당하는 A3~A5, A7, A8의 간략 정보가 보여진다. [그림 4-1]의 'Specific Query'에서 "윌트디즈니"를 선택하면, [그림 4-2]와 같이 'Current Query'에는 "20Y, 윌트디즈니"를

포함하고 있는 애니메이션의 개수인 "5작품(들)"이 표시된다.

[그림 4-2] [그림 4-1]의 Specific Query에서 "윌트디즈니"를 선택한 경우



'General Query'에는 "20Y, 윌트디즈니"의 상위 정점들이 개수인 "질의 수 : 2"와 질의 및 해당 작품들인 "200Y(3작품들)"과 "윌트디즈니(3작품들)"이 표시된다. 'Specific Query'에는 "질의 수 : 2"와 각 질의 및 해당 작품인 "3D(2작품들)"과 "어드벤처, 스티브 트랜버스(1작품)"이 표시된다. 'Results of Current Query'에는 'Current Query' 창의 결과인 A3, A4, A5의 간략 정보가 보여진다. 사용자가 브라우저에서 선택을 할 때마다 'History of Navigation'에서는 그 경로를 보여주어 앞 검색의 어느 시점으로도 바로 갈 수 있게 해준다.

5. 결론

현재 웹 상에서 모든 애니메이션의 검색 서비스를 제공하는 곳은 없으며, 일부 상업적인 목적을 가진 사이트에서 한정된 항목들만을 정리해놓고 검색할 수 있게 해주고 있다. 이러한 기존 사이트들의 검색 서비스를 이용하기 위해서는 사용자가 애니메이션 분류나 제작사와 같은 기본 지식을 가지고 있어서 직접 키워드를 타이핑해야 하거나, 제공하는 분류 기준에 따라 검색 기준을 바꿔가면서 여러 번에 걸쳐 검색해야 한다는 불편함이 있다. 본 논문에서는 애니메이션의 기본 정보를 래티스로 구성하고, 래티스를 적절히 항해할 수 있는 브라우저를 제안함으로써 사용자가 보다 애니메이션을 손쉽게 빠르게 검색할 수 있도록 하였다. 이 방법을 이용하면 용어들의 조합으로써 상세한 질의를 표현할 수 없는 일반 사용자들도 직접 키워드를 입력하지 않고 General Query나 Specific Query로 항해하면서 여러 키워드들의 조합으로 이루어진 질의를 만들 수 있어서 보다 손쉽게 원하는 결과를 얻을 수 있다. 그리고 여러 분류 기준에 해당하는 키워드의 조합을 한눈에 볼 수 있어서 기준을 달리하기며 여러 번 검색할 필요가 없으므로 보다 빨리 원하는 결과를 얻을 수 있다. 향후 연구로 시스템 구현을 통하여 본 논문에서 제안하는 시스템의 테스트를 실시할 것이다.

참고문헌

- Bestanime, <http://bestanime.co.kr/>
- Robert Godin, Jan Gecsel, and Clause Pichet, Design of a Browsing Interface for Information Retrieval, Proceedings of the 12th International ACM SIGIR Conference On Research & Development in Information Retrieval, 1989
- Salton, Gerard, Automatic Information Organization and Retrieval, McGrawHill, New York, 1968
- 유기영, 유관자 외, 이산치수학, 정익사, 1989

본 연구는 전북대학교 산업디자인 개발연구소의 지원을 받아 연구되었습니다.