

A Dynamic Associative Memory for Distant Reading

Ekaterina Kamlovskaya, Christoph Schommer, Joshgun Sirajzade
 ILIAS Lab, CSC Department, University Luxembourg

The talk is about the application of a *Dynamic Associative Memory*, a self-adapting organism able to expand and decrease itself depending on the data nutrition it gets, to historical texts. Originally introduced in (1) and further elaborated in (2,4,5), a dynamic associative memory is seen as a *collection of symbolic neural cells*, where each *cell* represents a *term* (for example, a word, a named entity, or an expression). The cells themselves share their own activation value; they become adhesive through a weighted connectivity which increases step-by-step if the terms occur together and gradually decrease if the terms do not, respectively. In this sense, the more often terms occur together, the stronger their connectivity will evolve (and vice versa). In order to avoid a too strong proliferation as well as decrease (in the sense of keeping the associative memory in a reasonable size), a variety of artificial mechanisms have been realized (for example, the concepts of 'fever' (too many connections) and 'cooling' (too little connections). If existing symbolic cells represent an identical term, then these cells as well as the associated memory structures they belong to are merged. Supporting pillars are a spreading activation of term stimulations and a *competitive learning* strategy in case of word ambiguity. In addition, the skeleton of the artificial organism is considered its centerpiece representing this part of the memory, which is seen as steady and stable over time.

In the context of historical texts, we apply dynamic associative memories to a corpus of 69 full-text autobiographies authored by people identifying as Australian Aboriginal or Torres Strait Islanders and published between 1950-2017. The research concern is manifold: on the one side, we are interested in receiving a *map of actors* at a specific time; on the other side, we raise the question of whether evolving skeletons (which are related to actors in a broader sense) can be interpreted as a topic or as a semantic structure of the text. In our simplest model, for example, in the book "Am I Black Enough for You" by Anita Heiss (3) one can find n-grams consisting of nouns and verbs with a flexible number of words between them within a sentence, such as {story}-{is}-{journey}, {country}-{is}-{New South Wales}, {spirit}-{belongs}-{ancestors}, etc. As

in the whole text of the book these words occur together multiple times, the connection between them is getting stronger. The connectivity weight hereby refers to the conditional probability of the cells. We build this n-grams of nouns combined by verbs by using NLTK POS Tagger and Named Entity Recognition. Stopwords are filtered out whereas nouns and named entities (including expressions) are kept; the results are associative structures that externalize a given content. We argue that these cells can describe the semantic skeleton of the book.

Key words (ACM Classification):

C.1.3 Adaptable Architectures; I.2.0 Cognitive Simulation; I.2.6 Connectionism and Neural Nets; I.2.7 Text Analysis.

References

- Christoph Schommer. Incremental Discovery of Association Rules with Dynamic Neural Cells. *Proceedings Workshop on Symbolic Networks, European Conference on Artificial Intelligence (ECAI)*, Valencia, Spain (2004).
- Claudine Brucks, Michael Hilker, Cynthia Wagner, Ralph Weires, Christoph Schommer. *Symbolic Computing with Incremental Mind-maps to Manage and Mine Data Streams - Some Applications. Proceedings 4th International Workshop on Neural-Symbolic Learning and Reasoning (NeSy'08)*. European Conference on Artificial Intelligence (ECAI), Patras, Greece (2008).
- Heiss, Anita, (2012). *Am I black enough for you?*. North Sydney, N.S.W : Bantam
- Jayanta Poray, Christoph Schommer. *Managing Conversational Streams with Explorative Mind-Maps. Proceedings ACS/IEEE International Conference on Computer Systems and Applications (AICCSA)*, Hammamet, Tunisia (2010).
- Jayanta Poray, Christoph Schommer. Operations on Conversational Mind-Graphs. *Proceedings 4th International Conference on Agents and Artificial Intelligence (ICAART)*, Vilamoura, Portugal. SciTePress (2012).

원거리 읽기(Distant Reading)를 위한 동적 연상 메모리

Ekaterina Kamlovskaya, Christoph Schommer, Joshgun Sirajzade
 룩셈부르크 대학 CSC 부서 ILIAS 연구실

이 강연은, 자기 적응 능력이 있는 유기체는 그가 취한 데이터라는 영양에 따라 스스로를 역사 텍스트로 확장하고 감소시킬 수 있는 동적 연상 메모리의 적용에 관한 강연이다. 처음 참고문헌 (1)에서 소개되고 참고문헌 (2,4,5)에서 상세히 설명된 동적 연상 메모리는 기호적 신경 셀의 집합으로 여겨지며, 여기서 각 셀은 용어를 표현한다(예를 들면, 단어, 개체명 또는 표현식 등). 셀들은 자체 활성화 값을 그들 스스로 공유하며, 가중 연결성을 통해 고착된다. 연결성은 해당 용어들이 함께 나타나는 경우에는 단계적으로 증가하고, 그렇지 않은 경우에는 점차적으로 감소한다. 이러한 의미에서 용어들이 함께 그리고 보다 자주 나타날수록 연결성은 더 강력하게 진화할 것이며, 그 역도 성립한다. 적절한 크기의 연상 메모리를 유지한다는 관점에서 너무 강하게 확산되거나 감소되는 것을 방지하기 위해, 다양한 인공 메커니즘을 구현시켜 왔다(예를 들면, '열'의 개념(연결이 개념이 너무 많음) 및 '냉각'의 개념(연결된 개념이 너무 적음)). 기존의 기호적 셀들이 동일한 용어를 나타내는 경우, 이들 셀 및 이들이 속한 관련 메모리 구조는 하나로 병합된다. 단어가 모호한 경우, 지지 기둥(supporting pillar)은 용어 자극의 활성화를 확산시키는 역할을 하며, 경쟁 학습 전략으로서의 역할을 한다. 또한, 인공 생명체의 골격은 기억의 이 부분을 표현하는 핵심 요소도 간주되며, 이는 시간이 흐르면서 지속적이고 안정적인 것으로 여겨진다.

우리는 역사적인 기록의 맥락에서 호주 원주민 또는 토레스 해협 제도 주민으로 신원이 밝혀진 사람들이 작성하고 1950~2017년 사이에 출판된 69권의 전체 텍스트 자서전의 언어 자료(corpus)에 동적 연상 메모리를 적용하였다. 특정한 시간의 행위자들의 맵(map of actors)을 받는 것에 관심을 두는 한편, 다른 한편으로 진화하는 골격(더 넓은 의미에서 배우들과 관련)이 주제 또는 본문의 의미론적 구조로서 해석될 수 있는지 여부에 대한 의문을 제기하는 등, 연구에 대한 우리의 관심사는 다양하다. 예를 들어, 우리의 가장 단순한 모델을 이용해도, Anita Heiss(3)가 저술한 "Am I Black Enough for You"라는 책에서, {story}-{is}-{journey}, {country}-{is}-{New South Wales}, {spirit}-{belongs}-{ancestors} 등과 같이, 한 문장 내에서 유연한 수의 단어를 사용해 명사와 동사로 구성된 n-그램을 찾을 수 있다. 책의 전문에서와 같이, 이들 단어가 여러 번 함께 나올수록, 그들 사이의 연결성은 점점 더 강해진다. 여기서 연결성 가중치는 셀의 조건부 확률을 말한다. 우리는 NLTK POS 태거(NLTK POS Tagger) 및 개체명 인식(Named Entity Recognition)을 사용하여 동사와 결합된 명사의 n-그램을 만든다. 제외어는 필터링되어 없어지지만 명사와 개체명(표현식 포

함)은 없어지지 않고 유지되며, 결과는 주어진 내용을 최적화하는 연상 구조이다. 우리는 이 셀들을 통해 책의 의미론적 골격을 기술할 수 있다고 주장한다.

키워드 (미국컴퓨터학회(ACM) 분류):

C.1.3 적용 아키텍처, I.2.0 인지적 시뮬레이션, I.2.6 연결주의 및 신경망, I.2.7 텍스트 분석.

참조문헌

- Christoph Schommer. 동적 신경 셀을 이용한 연관 규칙의 점진적 발견 기호적 네트워크에 관한 워크샵 회보 유럽인공지능회의(ECAI), 스페인 발렌시아(2004).
- Claudine Brucks, Michael Hilker, Cynthia Wagner, Ralph Weires, Christoph Schommer. 데이터 스트림의 관리 및 채광을 위한 점진적 마인드맵을 이용한 기호적 컴퓨팅-일부 애플리케이션. 신경-기호 학습 및 추론에 관한 제4차 국제 워크숍 회보(NeSy'08) 유럽인공지능회의(ECAI), 그리스 파트라스(2008).
- Heiss, Anita (2012) *너가 보기에 내가 그렇게 겁니?* 뉴사우스웨일즈 북부 시드니 : Bantam
- Jayanta Poray, Christoph Schommer. *예비적 마인드맵을 이용한 대화형 스트림의 관리. 컴퓨터 시스템 및 애플리케이션에 관한 ACS/IEEE 국제학술회의(AICCSA) 회보*, 튀니지 하마멧(2010)
- Jayanta Poray, Christoph Schommer. *대화형 마인드 그래프에 관한 작업 제4차 에이전트 및 인공 지능 국제학술회의(ICAART) 회보*, 포르투갈, 빌라모우라. Sc.ePress(2012).