

AI를 통한 글쓰기와 작가의 운명 —「コンピュータが小説を書く日」을 중심으로*—

이 정 엽**

요 약

이 논문은 컴퓨터 인공지능을 통한 창의적인 글쓰기의 사례 분석을 통해 최근 제기되고 있는 작가의 운명과 존재론에 대한 시론적인 접근을하고자 한다. 일본 나고야대학교 사토 사토시 교수의 연구실에서 제작한 인공지능이 창작한 소설 「컴퓨터가 소설을 쓰는 날(コンピュータが小説を書く日)」을 중심으로 이 소설에 사용된 자연어 처리의 알고리즘을 분석하고, 그 소설적 성취를 평가하고자 했다. 「컴퓨터가 소설을 쓰는 날」은 기계 학습을 위해 기존 소설의 문장을 모방하고 치환하거나, 짧은 회화로 이루어진 대화체 소설을 기존 소설의 여러 번을 통해 병치시키는 방법으로 소설을 창작해 왔다. 그러나 이러한 방법론으로 컴퓨터에게 완전히 창의적인 소설을 스스로 쓰는 방법을 학습시킬 수는 없었으며, 이는 인공지능을 통한 글쓰기가 가지는 일정한 한계를 보여준다.

주제어: 인공지능, 소설 창작, 스토리 자동 생성, AI, 컴퓨터가 소설을 쓰는 날

* 본 연구는 순천향대학교 학술연구비 지원으로 수행하였음. (20171101)

** 순천향대학교 한국문화콘텐츠학과 조교수

목차

1. 서론
2. AI를 통한 글쓰기와 그 전개과정
3. 「컴퓨터가 소설을 쓰는 날」의 알고리즘 분석
4. 결론: AI 작가의 등장과 존재론

1. 서론

“컴퓨터는 과연 인간처럼 글을 쓸 수 있을 것인가?” 이러한 질문을 던지는 것이 상당히 우스운 이야기가 되었던 시기도 있었다. 컴퓨터는 그저 인간이 만든 프로그램을 순서대로 처리하는 단말기에 불과했던 상황에서는 이러한 질문은 특별한 근거 없이 상상의 차원에서 비약이 들어간 질문에 불과했기 때문이다. 그러나 해가 갈수록 컴퓨터의 처리 속도가 빨라지고, 컴퓨터가 상황을 인식하고 학습하는 방식이 바뀌게 되면서 컴퓨터도 일정한 제약이 존재하기는 하지만 특정한 패턴에 맞추어 글을 쓰는 시도들이 늘어나게 되었다. 최근에는 자연어 처리(natural language processing) 기능과 머신러닝(machine learning) 혹은 딥 러닝(deep learning)에 기반한 새로운 프로그램들이 속속들이 개발되고 있으며, 이들이 적용되는 분야는 산업 일반과 유통, 데이터를 넘어 인간 고유의 영역으로 알려져 있는 창작의 분야까지 넘보고 있는 상황이 되어버렸다.

이제 컴퓨터는 주어진 데이터만 조건에 맞으면 프로야구나 주식, 날씨 예보 등에 관한 한 인간보다 더 빠르게 기사를 작성할 수 있다. 이른바 ‘로봇 저널리즘(robot journalism)’¹⁾이라는 분야는 이

1) 기사를 자동 생성하는 로봇 저널리즘에 관한 국내의 연구로는 김동환, 이준환, 「로봇 저널리즘 : 알고리즘을 통한 스포츠 기사 자동 생성에 관한 연구」, 한국

미 국내에도 많은 언론사들이 활용하고 있으며, 앞으로는 많은 기자들의 일자리를 실질적으로 위협할 것으로 판단된다. 프로야구 기사나 주식 기사 등에서 특별한 창의성을 요구하는 문장이 필요 없고, 대부분 동일한 패턴의 문장이 반복되기 때문에 이를 컴퓨터에게 학습시켜 데이터만 충분히 주어진다면 상황에 맞는 문장들을 조합하여 빠르게 기사를 생성할 수 있다. 때문에 비예술적인 분야에서는 일찌감치 자연어처리와 인공지능을 바탕으로 한 문장 생성에 관심을 두어 왔다. 그러나 소설이나 시와 같은 문예적인 작품의 창작은 여전히 인간의 고유한 분야인 것처럼 인식되고 있는 것도 사실이다. 그러나 의외로 컴퓨터를 통한 문예 작품의 자동 창작은 1970년대 후반 연구되었으며, 그 흐름은 ‘드라마티카 프로’²⁾나 국내에서 개발된 ‘스토리 헬퍼’³⁾ 같은 스토리 저작 도구와 ‘테일스핀(Tailspin)’이나 ‘민스트렐(Minstrel)’ 같은 스토리 자동 생성 프로그램으로 나뉘 볼 수 있다.⁴⁾

언론학보 제59권 5호, 2015.를 참고할 것.

2) <http://dramatica.com>

3) <http://www.storyhelper.co.kr/>

4) 스토리 창작 도구와 자동 생성 프로그램의 관계에 관한 국내의 연구는 그리 많지 않다. 참고할만한 연구 목록은 다음과 같다.

이인화, 『스토리텔링 진화론: 창작의 원리에서 도구까지, 위대한 이야기는 어떻게 만들어지는가?』, 해냄, 2014.

권호창, 『디지털 스토리텔링에 있어 스토리 생성 시스템 연구: 스토리 엔진 모델링을 중심으로』, 한국예술종합학교 예술전문사 학위논문, 2010.

류철균, 서성은, 『디지털 서사 창작 도구의 서사 알고리즘 연구: 드라마티카 프로를 중심으로』, 『현대소설연구』 제38권, 2008.

류철균, 윤혜영, 『디지털 서사 창작도구의 CBR 모델 비교 연구 : 〈민스트렐〉과 〈스토리헬퍼〉를 중심으로』, 『디지털컨텐츠학회논문지』 제13권 제2호, 한국디지털컨텐츠학회, 2012.

윤현정, 『주역을 활용한 한국형 스토리텔링 창작 지원 도구 개발 방향 연구』, 『디지털 스토리텔링 연구』 제5권, 2010.

이찬욱, 이채영, 『영화·애니메이션 스토리텔링 기획·창작지원시스템 연구』, 『인문콘텐츠』 제19호, 인문콘텐츠학회, 2010.

드라마티카 프로는 창조적인 글쓰기를 하는 작가들을 대상으로 그들의 작품이 더 창의적이 될 수 있도록 작품의 기초와 캐릭터, 주제, 플롯, 장르 등에 관한 질문을 작가에게 계속해서 던지면서 작품이 다른 작품과의 유사성을 피해 상투적이지 않은 작품을 유도하는 것을 목표로 삼고 있는 프로그램이다. 스토리 자동 생성 프로그램을 연구했던 이인화와 이화여자대학교 연구팀이 엔씨소프트와 협력하여 개발한 ‘스토리 헬퍼’는 드라마티카 프로의 단점을 개선한 모티프 조합 방식을 사용했다. 이러한 스토리 저작 도구는 ‘사례 기반 추론 모델(case-based reasoning model)’에 의거하여 귀납적인 방법을 통해 과거의 작품과 창작 중인 새로운 작품을 고안해 낸다. 반면 서사 자동 창작 도구는 기존의 이야기들을 프로그래밍 함수로 된 스토리 문법으로 체계화하고 이를 바탕으로 문장 또는 텍스트 단위의 이야기를 자동 생성한다. 스토리 저작 도구와 스토리 자동 생성 프로그램은 서사의 창작이라는 동일한 연구 주제에 대해 문학과 공학이라는 전혀 다른 학문적 배경에서 접근하고 있는 것이다.⁵⁾ 본 논문에서 살펴보고자 하는 『컴퓨터가小説을書く日』은 스토리 자동 생성 프로그램이 창작한 하나의 예시 소설에 해당된다.

일본 나고야대학교에서 자연어 처리와 인공지능을 연구하고 있는 사토 사토시(佐藤理史) 교수 연구실은 2015년 일본 경제신문에서 주관하는 제3회 호시이 신이치(星新一)⁶⁾ 문학상에 컴퓨터를 통한

혹은 이 분야에 관한 역사적인 맥락과 전개 과정을 알고 싶으면 Noah Wardrip-Pruin, *Expressive Processing: Digital Fictions, Computer Games, and Software Studies*, The MIT Press, 2012.를 참고할 것.

5) 류철균, 윤혜영, 『디지털 서사 창작도구의 CBR 모델 비교 연구 : 〈민스트럴〉과 〈스토리헬퍼〉를 중심으로』, 『디지털컨텐츠학회논문지』 제13권 제2호, 한국디지털컨텐츠학회, 2012, p.214.

6) 호시이 신이치는 일본의 유명 SF 작가로 ‘쇼트 쇼트(초단편소설)’라는 새로운 분야를 개척했다. 1926년 생으로 도쿄에서 태어나 도쿄대학을 졸업한 그는 1957년 SF동인지 『우주진』에 발표한 ‘세키스토라가 『보석지』에 연재되면서 작

인공지능으로 쓴 소설 작품을 공모했다. 그 중 한 편을 인터넷을 통해 공개했고, 호의적인 반응을 얻은 사토 교수팀은 이듬해 제4회 호시이 신이치 문학상에도 인공지능으로 쓴 소설 작품을 공모했다. 『컴퓨터가 소설을 쓰는 날(コンピュータが小説を書く日)』이라는 A4 3장 정도의 이 짧은 단편 소설은 국내의 여러 언론에도 소개되고 전문이 공개되면서 관심을 모았다. 지금까지 개발된 대부분의 스토리 자동 생성 프로그램은 미국을 비롯한 서구에서 진행된 것이다. 그러나 사토 교수 팀에서 만든 일본어 소설 자동 생성 프로그램은 일본어와 한국어의 어순이 거의 동일하고 많은 부분에서 쉽게 상대 언어로 치환과 번역이 가능하다는 점에서 주목할 만하다.

컴퓨터를 통한 인공지능이 소설을 쓸 수 있게 되었다는 사실은 소설 쓰기라는 다분히 인간적인 예술 활동에 대한 일종의 도전처럼 여겨졌기 때문에 일반 대중에게도 충격적으로 다가왔다. 이러한 충격은 컴퓨터에 의해 대체되지 않을 것으로 생각했던 작가의 고유성(固有性)과 창조적인 활동이 기계적인 알고리즘으로 대체될 수 있다는 점에서 오는 것이라고 할 수 있다. 인간과 컴퓨터는 서로 대척적인 지점에 놓인 것처럼 생각하기 쉽지만, 컴퓨터를 통해 서사 창작을 탐구하는 것이 작가의 고유성을 부정하는 행위라고 보기는 어렵다. 『컴퓨터가 소설을 쓰는 날』을 개발한 사토 사토시 교수도 다음과 같이 밝힌 바 있다.

인공 지능에 의한 자동 창작 연구는 탐구할만한 가치가 있

가로 데뷔했다. 1961년 처녀 단편집 〈인조미인〉으로 나오키상을 수상했으며 1968년 〈망상은행〉으로 일본추리작가협회상을 수상했다. 지금까지 1000편이 넘는 작품을 발표했다. 극도로 짧은 폰트 같은 쇼트 쇼트 장르는 전체 길이가 짧기 때문에 컴퓨터에게 창작 연습을 시킬 때 모방의 좋은 전범이 된다. 일본경제신문사에서 호시이 신이치 문학상을 제정하면서 컴퓨터 인공지능에 기반한 작품도 공모의 대상이 된다고 지칭한 것은 호시이 신이치 문학이 가진 특징을 반영한 처사라고 할 수 있다.

다. 왜냐하면 인간과는 다른 창작 방법을 통하여 컴퓨터는 예기치 못한 새로운 예술적 가치를 창출할 수 있기 때문이다. 또한 기계적인 창작 기술은 인간의 창작의 폭을 넓히는 데 도움이 될 수 있다.⁷⁾

기본적으로 컴퓨터가 소설을 쓰기 위해서는 인간이 어떻게 소설을 써왔는지 그 순서와 과정을 컴퓨터에게 이해시키는 것이 필요하다. 당연히 소설 역시 인간이 오랫동안 써왔던 다양한 글쓰기의 일부분인 한 형태이기 때문에, 당연한 것처럼 여겨져 온 소설 쓰기 방식을 컴퓨터에서 절차적인 형태(procedural process)⁸⁾로 구현해 낼 필요가 있다.

여기서 중요한 것은 컴퓨터가 소설을 쓰는 전 과정을 모두 다 인간적인 관점에서 이해하고 전체적인 상황을 컨트롤 할 필요가 없다는 점이다. 컴퓨터는 부분적으로 소설 속의 문장을 이어나가는 패턴을 학습하고, 이를 개발자가 마련한 알고리즘에 기반하여 텍스트의 형태로 구현해내지만 하면 된다. 인공지능이 만들어 낸 문장과 문장, 혹은 문단과 문단 사이의 의미를 연결지어주는 맥락(context)은 모두 컴퓨터의 자율 의지에 의해 작성된 것이 아닐 수도 있으며, 개발자의 인위적인 조정에 의해 취사선택된 것일 수 있다. 그렇다고 하더라도 독자의 입장에서는 인공지능이 완성한 한 편의 소설을 완결된 작품으로 읽으려는 맥락 중심주의적인 독해 방식 혹은 습관 때

7) 緒方健人, 佐藤理史, 駒谷和範, 『模倣と置換に基づく超短編小説の自動生成』, 『The 28th Annual Conference of the Japanese Society for Artificial Intelligence』, 2014, p.1.

8) 절차성(procedurality)는 컴퓨터 매체가 가진 본질적인 속성 중의 하나로, 순서와 절차에 의해 주어진 프로그램을 실행하는 능력을 의미한다. 자넷 머레이(Janet Murray)는 이러한 절차성을 디지털 미디어가 갖추어야 할 본질적인 속성 4가지(절차성, 참여성, 공간성, 백과사전적 구조) 중 하나로 들고 있다. Janet Murray, *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*, The MIT Press, 2003, pp.71-78.

문에, 실제 이 작품을 쓴 컴퓨터가 모든 의지와 능력을 가지고 소설을 창작한 것처럼 독해할 수 있다는 것이다. 그러나 실제로 컴퓨터는 소설을 창작하는 방법과 문장을 쓰는 방식에 대한 모든 이해를 갖춘 것처럼 스스로의 결점을 감추고만 있으면 되는 것이다. 왜냐하면 소설 텍스트에는 표층적인 문장만 드러나지 실제 그것을 창작한 기저에 깔린 심층적인 알고리즘(algorithm)은 겉으로 전혀 드러나지 않기 때문이다.

그렇다고 해서 이 논문이 컴퓨터가 만들어 낸 인공지능 소설의 가치를 부정하는 것은 아니다. 컴퓨터가 창작한 소설의 알고리즘을 분석해 나가다보면 인간이 소설을 쓰는 과정을 부분적으로나마 질차적으로 규명해 낼 수 있으며, 이러한 부분은 실제 소설 창작에도 많은 시사점을 제공해 줄 수 있기 때문이다. 사토 교수의 언급처럼 컴퓨터를 통한 창작 방식을 통해 우연치 않게 새로운 예술적 가치를 창출할 수도 있기 때문이다. 이 논문은 소설을 창작하는 각각의 개인과는 별개로 존재하는 서사 창작의 원리를 디지털 시대에 맞게 재해석하려고 한다. 실제로 컴퓨터 혹은 인공지능이 쓴 소설의 성취를 분석하기 위해서는 표층적으로 드러난 작품의 전문만을 보아서 이를 정확하게 분석할 수 없다. 컴퓨터에 작성된 소설 창작의 알고리즘은 기본적으로 이러한 소설을 무한히 다른 패턴으로 반복적으로 창작해 낼 수 있는 기저에 깔린 메커닉(mechanic)⁹⁾이기 때문이다. 때문에 이 논문은 컴퓨터의 인공지능을 활용한 사토 사토시

9) 디지털 미디어에서 메커닉은 이러한 미디어를 가동하기 위한 눈에 보이지 않는 심층적인 요소를 의미한다. 주로 프로그래밍에 의해 만들어진 알고리즘이나 규칙 등을 의미한다. 이러한 메커닉은 눈에 보이거나 들리는 시청각적인 요소인 다이내믹(dynamic)과 결합하여 최종적으로 사용자의 감성을 자극하는 미학적 요소로 치환된다. 이렇듯 메커닉-다이내믹-미학적 요소로 이어지는 순환의 과정은 Marc LeBlanc, "Tools for Creating Dramatic Game Dynamics", *The Game Design Reader: Rules of Play Anthology*, The MIT Press, 2005, pp.438-459.를 참고할 것.

교수팀의 스토리 자동생성 도구의 알고리즘을 분석하고, 이를 기반으로 컴퓨터가 만든 소설을 바탕으로 작가의 존재론에 대해 다시 한번 고찰해 보는 것을 목표로 한다.

최근 언론과 미디어를 중심으로 많이 회자되고 있는 ‘4차 산업혁명’이라는 전환기에는 다양한 ICT 테크놀로지가 필요하다. 인공지능은 빅데이터, 사물인터넷, 클라우드, 모바일 기술 등과 함께 생산성의 혁신을 가져올 핵심 테크놀로지로 소개된다. 실제로 이러한 테크놀로지가 생산성을 혁신하고 삶의 방식을 획기적으로 전환할지에 대한 것은 이 논문의 관심사가 아니다. 여기에서는 이러한 거시적인 관점에서 인공지능을 논하기보다 보다 미시적인 관점에서 인공지능을 통한 글쓰기 방식을 분석하고, 그 알고리즘에 내재된 논리를 서사적으로 해명하는 것을 목표로 한다.

2. AI를 통한 소설 쓰기 방식과 그 구현 과정

1) 모방과 치환을 통한 문장 작성

사토 사토시 교수는 처음부터 컴퓨터 인공지능을 바탕으로 소설을 자동 창작하는 것에 관심을 기울이지는 않았다. 그러나 그는 2012년 호시이 신이치 문학상이 제정된 후 이에 관심을 가졌던 하코다테 미래대학의 마츠바라 히토시(松原仁) 교수 팀으로부터 인공지능에 기반한 소설을 같이 쓰자는 제안을 받게 된 이후로 이 주제에 처음으로 관심을 쏟게 된다.¹⁰⁾

두 교수 모두 일본어를 기반으로 한 자연어 처리를 공통의 주제

10) 佐藤理史, 『コンピュータが小説を書く日: AI作家に「賞」は取れるか』, 日本経済新聞出版社, 2016, p. 8.

로 삼고 있었기 때문에, 각각의 연구를 발전시키면서 의논해 나가면 좋은 작품을 쓸 수 있다는 것이 마츠바라 교수의 제안이었던 것이다. 초기에 사토 교수와 마츠바라 교수는 공통의 주제를 가지고 학술적인 논의를 진행해 나가지만, 컴퓨터에게 소설을 구현하게 만드는 최초의 설정 방식과 물리적인 거리 등으로 인해 별도로 연구를 진행해 나가게 된다.

사토 교수는 컴퓨터가 처음부터 긴 소설을 창작하는 것이 힘들다고 판단하고 트위터의 한 문장 제약 길이인 140자 정도의 짧은 문단을 컴퓨터를 통해 자동 생성하는 것을 목표로 삼았다. 2014년 일본인공지능학회에서 발표한 논문 “모방과 치환에 기초한 초단편소설의 자동생성 (ACJSA 2014)”에서 그는 호시이 신이치 작품의 여러 문장을 발췌하여 그 문장을 절 단위로 섞어 컴퓨터가 새로운 논문을 만드는 방식을 고안해 낸다.

(이 논문에서 소설을 쓰는 방식은) 기존의 작품에서 문장을 추출하여 그 일부를 다른 작품의 일부분으로 대체하는 방법이다. 이 대체 단위로 “술어(術語)을 중심으로 한 정리”를 나타내는 절(節)이라는 언어 단위를 채택한다. 또한 대체 대상을 문장의 골격을 제외한 종속절에 한정하여 바꿀 문장이 크게 파괴되는 것을 억제한다. 앞서 말한 에스카 유(江坂遊)의 방법은 단어 단위로 조합의 조합보다는 발상을 얻은 인간이 창작을 하는 방법인데, 이 방법은 말하자면 절을 결합하여 그것을 그대로 문장으로 채택 하자는 방법으로 간주될 수도 있다.¹¹⁾

컴퓨터는 아무 것도 학습되지 않은 상태에서 ‘무(無)’의 관점에서 소설 쓰기를 시작할 수는 없다. 이에 사토 교수는 호시이 신이치가

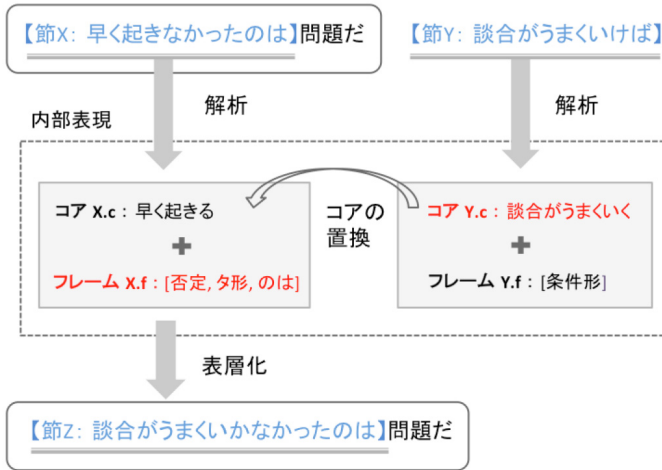
11) 緒方健人, 佐藤理史, 駒谷和範, 『模倣と置換に基づく超短編小説の自動生成』, 『The 28th Annual Conference of the Japanese Society for Artificial Intelligence』, 2014, p.1.

고안한 ‘쇼트쇼트(초단편소설)’ 장르를 전범으로 하여 이 문장들을 절(節) 단위로 재조합하는 방식으로 컴퓨터가 쓰는 문장 방식을 고안해낸다. 호시이 신이치가 쓴 문장이 가진 신뢰를 바탕으로 어느 정도 의미를 가지고 있다고 생각되는 절 단위를 다른 호시이 신이치 문장에서 가져와서 종속절을 변화시키는 방법으로 문장을 재조합하는 것이다.

예술 작품은 다른 사람이 시도하지 않은 새로움과 독창성을 갖춰야 하지만, 사토 교수팀의 첫 출발은 다른 소설가의 문장을 모방하여, 이를 같은 소설가의 다른 문장과 결합하여 종속절을 약간 변형하는 정도의 가공을 통해 완전한 모방을 피하는 방식으로 첫 창작을 시작하는 것이다.

사토 교수팀은 전체의 문장을 그 내용에 해당하는 코어(core) 부분과 그 문장의 뉘앙스, 형식 등을 담은 프레임(frame) 부분으로 나눈다. X라는 작품에서 따온 문장 ‘일찍 일어나지 않는 것이 문제다’에서 프레임만 남기고 코어에 해당하는 ‘일찍 일어나다’라는 내용은 삭제한 후, Y라는 작품에서 따온 종속절 ‘대화를 이어가는 것은’이라는 문장에서 코어에 해당하는 내용인 ‘대화를 이어가는 것’을 서로 치환한다.¹²⁾ 이러면 ‘대화를 이어가는 것이 문제다’라는 문장이 되는데 의미상으로 부정적인 뉘앙스를 풍기게 되면 문장이 더 자연스럽게 때문에 프레임을 부정적인 뉘앙스로 변형시켜 ‘대화를 이어 나가지 않는 것이 문제다’라는 문장으로 새롭게 변형시킨다. 사토 교수팀은 이러한 프레임의 2차적인 변형을 ‘표층화(表層化)’라 부른다. 실제 독자가 만나게 되는 가장 표층적인 형태의 문장 변형이기 때문에 이러한 용어를 사용한 것으로 보인다.

12) Ibid., p.3.



〈그림 1〉 종속절을 치환하는 문장 작성 개념도

이러한 과정을 거쳐 ‘대화를 이어나가지 않는 것이 문제다’라는 그럴듯한 문장을 하나 창작해내게 되는 셈이다. 아무런 연관도 없는 두 문장이 쉽게 결합할 수 있는 이유는 이러한 작품이 140자 정도의 짧은 초단편소설이어서 작품의 서두에 형성된 맥락이 전혀 없어서 어떠한 내용도 다 수용할 수 있기 때문이다. 비디오 게임의 초기 연구에 기여했던 브렌다 로렐(Brenda Laurel)은 어떤 작품이 처음에 시작할 때에는 모든 문장이 가능성의 영역에 놓이지만, 작품이 진행될수록 모든 문장이 다 가능한 것은 아니고 앞서 형성한 맥락에 따라 개연성이 존재하는 문장들만 이어져 나갈 수 있다는 사실을 밝힌 바 있다.¹³⁾ 사토 교수 팀은 이러한 점을 의식하여 의도적으로 맥락적인 부분이 굉장히 약한 고리를 가져도 되는 초단편소설만을 대상으로 하여 초기의 문장 생성을 시도한 것으로 판단된다.

13) Brenda Laurel, *Computers as Theatre*, Addison-Wesley, 1991, p.68.

2) 회화를 중심으로 하는 초단편 소설의 창작

초단편소설을 완성하는 두 번째 방법은 기존 소설에 나와 있는 회화를 중심으로 이들을 재조합하는 것이다. 사토 교수 팀은 마츠바라 교수의 제안을 받아 이 분야의 연구를 시작하면서 일본어 자연어 처리 연구에 있어서 문장의 자동생성 분야 연구 수준이 특히 낮다고 판단했다.

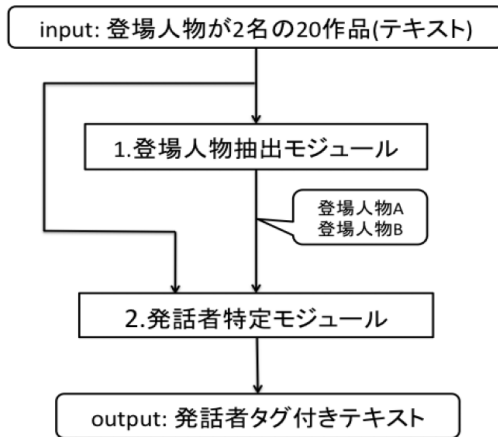
단편 자동 창작에 많은 기능의 자동화가 필요하다. 그 중에서 가장 큰 걸림돌이 될 것으로 예상되는 것은 문장 생성, 즉 “인간이 읽고 의미를 취할 수 있으며 위화감을 느끼지 않는 한 단락 이상의 텍스트를 만드는 것”이다. 일본어 문장생성 연구는 분석 연구에 비해 많이 늦어 아직까지도 문장을 자동으로 생성 도구조차 존재하지 않는다.

이러한 인식에 서서 우리가 만드는 소설의 길이를 140 자, 400 자, 1200 자와 단계적으로 길게 해 나가는 전략을 취한다. 첫 140 자는 트위터 소설이라고 트위터의 등장과 함께 출현 한 소설 형식이다. 본 논문에서는 140 자 이내의 초단편 소설을, 회화를 중심으로 조립하는 방법을 검토한다. 대화는 소설을 구성하는 중요한 요소 중 하나이며, 로버트 B 파커의 스펜서 시리즈처럼 대화가 매력이 되는 소설도 많이 존재한다.¹⁴⁾

우선 짧은 단락 하나 정도의 길이를 가진 회화를 중심으로 한 초단편소설을 창작하는데, 이번에도 문장 생성을 위해 호시이 신이치의 단편 2편을 그 전범으로 한다. 호시이 신이치의 『샘(泉)』이라는 단편에서 A와 B라는 2명의 인물이 나누는 회화 4문장을 발췌한 뒤, 이 회화에 A1B1A2B2라는 식의 태그를 붙인다. 대부분의 소설

14) 高木大生, 佐藤理史, 駒谷和範, 『會話を中心とした超短編小説の自動生成』, 『The 28th Annual Conference of the Japanese Society for Artificial Intelligence』, 2014. pp. 1-4.

에는 회화를 발화하는 대상을 제대로 지칭하지 않고 생략하는 경우가 많으니 이러한 각각의 발화가 어느 인물로부터 비롯된 것인지 정확히 기입해준다. 또 호시이 신이치의 또 다른 단편 『크고 작은 사고(小さくて大きな事故)』로부터 또 다시 2명의 다른 인물 C와 D가 나누는 회화 4문장을 발췌하여 이를 C1D1C2D2 형태로 태그를 붙인다. 이를 바탕으로 두 회화 중 일부를 서로 치환하여 인물을 재조합하는 방식으로 A1B1A2B2 중 동일 인물의 발화를 재편성, 즉 A1D1A2D2 또는 C1B1C2B2을 창조한다. 혹은 후반부의 재조합 A1B1A2B2 중 후 2 발화를 재편성하여 A1B1C2D2을 창조한다.



〈그림 2〉 발화자를 특정하는 시스템 구성도

사실 전혀 다른 두 소설로부터 동일 인물인 것처럼 생각하기 쉬운 2명의 인물을 동일인물인 것처럼 위장하여 서로의 문장을 바꿔치기 하는 방법은 일단 무엇보다도 표절에 해당된다. 그럼에도 불구하고 사토 교수 팀이 이러한 방법을 선택한 이유는 기계에 의한 소설 생성 먼저이 수준의 습작부터 시작할 필요가 있다고 판단했기 때문

이다. 이 후 발화의 일부를 변경하거나 여타의 문장을 삽입하거나하여 표절의 정도를 희석하는 것이 가능하다. 이렇게 조합하여 만들어진 회화가 비슷한 성격을 가진 인물인지 감지하기 위해 문장을 형태소별로 나누어 인식한다. 이렇게 추출된 형태소를 분석하여 유의미한 매칭이 이루어지면 회화가 새롭게 성립한다고 말할 수 있게 되는 것이다.

表 3: 登場人物抽出の正解率

登場人物の総数に対する正解率	42/46	91%
作品の総数に対する正解率	20/23	87%

〈그림 3〉 등장인물을 제대로 추출하는 확률

그러나 이러한 방법론 역시 다른 소설로부터 전범을 빌려왔다는 점에서 창조성에 관한 근본적인 한계를 가지고 있다. 표절이라는 비난을 피하기 위해 서로 다른 소설로부터 회화를 짜깁기하고 있을 뿐 대부분의 문장들이 컴퓨터가 직접적으로 창작한 문장은 아니라는 한계를 노출한다. 그렇게 다른 작품으로부터 회화를 발췌하다보니 회화들 사이의 맥락 또한 큰 손상을 입게 된다. 현재로서는 이렇게 창작된 작품은 두 소설 사이에서 임의로 발췌된 회화가 자아내는 약간의 의외성적인 측면에서 나오는 우연적인 미학 외에는 큰 성과가 없다고 할 것이다.

3) 플롯을 중심으로 하는 단편 소설의 창작

상황이 이러하다보니 사토 교수팀은 보다 근본적인 차원에서 긴 소설을 컴퓨터의 능동적인 측면을 강화한 소설을 쓸 필요성을 느끼

게 된다. 앞선 두 모델의 경우 간신히 전면 표질의 혐의만 벗었을 뿐, 가치 있는 경험을 전달하고자 하는 문학의 근본적인 목표에는 전혀 도달하지 못하고 있었다. 이에 사토 교수팀은 방법을 근원적으로 바꾸어 컴퓨터가 직접 플롯을 생성하고 인간이 문장을 쓰는 것을 보좌하는 방법을 취하고자 했다.

소설을 쓰기 위해서는 우선 플롯을 만들고 그것을 문장화하는 것이 필요하다. 여기서는 4작품 중 두 작품을 만든 ‘늑대 인간 지능 프로젝트’ 플롯 작성 부분을 자동화하여 얻어진 플롯을 인간이 문장화하는 방식을 사용했다. 나머지 두 작품은 만든 ‘우연히 인공 지능 프로젝트 작가가 되었어요.’ 프로젝트를 통해 인간이 플롯을 만들고 그것을 컴퓨터를 사용하여 문장화했다.¹⁵⁾

2017년에 사토 교수 팀이 세 번째로 시도한 방법은 인간과 컴퓨터가 서로 협력하여 한 쪽이 플롯을 만들면 한 쪽이 문장을 작성하는 방식이다. 반 정도 인공지능에 가까워졌기 때문에 인공지능이란 말을 패러디하여 인랑지능(人狼知能)이라 표현한 것으로 보인다. 여기에서 플롯은 인랑(人狼)이라는 게임의 로그(log) 데이터를 활용하여 여기에서 플레이어가 취한 행동을 플롯으로 지칭하는 것이다. 게임 내에서 플레이어 캐릭터가 취하는 행동은 몇 가지의 정량화된 패턴을 지니기 때문에, 이를 엄밀한 의미에서 플롯이라 말하기는 어려울 것이다. 우선 사건과 사건 사이의 인과관계가 드러나야 하는 플롯의 본질이 훼손되고, 플레이어의 의지에 따라 임의적으로 선택된 행동들이 왜 그러한 행동을 취했는지에 대한 의도적 분석 없이 명령어 형태로 로그파일에 실려 있기 때문이다.

15) 松山諒平, 佐藤理史, 松崎拓也 『人狼ログからの小説の自動生成』, 『言語処理學會 第23回年次大會 發表論文集』, 2017. 03, p.32.

```

1,talk,5,10,COMINGOUT Agent[10] MEDIUM
1,talk,6,3,DIVINED Agent[01] HUMAN
1,talk,7,8,VOTE Agent[10]
1,talk,8,5,VOTE Agent[02]
1,talk,9,4,VOTE Agent[10]
1,whisper,0,7,VOTE Agent[10]
1,whisper,1,6,Over
:
1,divine,3,9,HUMAN
1,guard,1,3,SEER
1,attackVote,6,1
1,attackVote,7,3
1,execute,2,VILLAGER
1,attack,3,false
:

```

图 1: 人狼ログの例

〈그림 4〉 인랑 로그 파일의 예시

사토 교수팀은 문학 작품에서 문장 형태로 나타나는 결과, 즉 표층화된 부분의 결과물에 너무 큰 관심을 쏟은 나머지 실제 문학 작품의 의미화에 매우 중요한 맥락의 지속적 생성을 컴퓨터가 인식하고 자동으로 만드는 부분을 구현하는 데에 실패한다. 인간이 만든 소설에서 지문과 대화를 발췌하여 섞고, 게임 로그로부터 추출한 원시적인 플롯 데이터를 바탕으로 인간과 컴퓨터가 서로 협력하여 플롯과 문장을 번갈아 작성하는 방식은 각 플롯 사이의 극적인 구조가 생략되어 있어 각각의 플롯을 극화하기 쉽지 않다는 문제가 발생한다.

모방과 치환을 통한 방법론에서 종속절의 치환을 할 때에도 종속절을 현재 문단에 맞게 수정하기 위해서는 다양한 형태의 접속사와 그 뉘앙스를 컴퓨터에게 학습시키고, 이를 현재의 문장 상황에 맞추는 것만 해도 일본어 동사 유형에 따라 7개의 모델링이 필요했다.¹⁶⁾

基本形	起きない
条件形	起きなければ
連用形	起きず
タ形	起きなかった
タ系条件形	起きなかつたら
テ形	起きずに
タリ形	起きなかつたり

〈그림 5〉 종속절 치환을 위한 7개의 유형

그러나 이러한 시도에도 불구하고 결국 사토 교수 팀은 이와 같은 로그 데이터를 활용한 플롯 생성 방식으로 소설을 쓰는 것이 결과적으로 실패했음을 논문을 통해 밝히고 있다. 기본적인 캐릭터의 설정을 모방과 치환을 통해 도출시킨 문장이 반영을 하지 못하는 것도 문제이지만, 일반적인 소설이 내재하고 있는 주제적인 차원을 접근을 전혀 시도하지 못한 점이 가장 큰 문제 요인이 된 셈이다.

더 심각한 문제는 생성된 소설에는 다음과 같은 요소가 부족하다는 점이다.

○ 캐릭터 설정: F 메구미(이 소설의 주인공)는 평소에는 기사로 일하고 있는 등의 간단한 설정이지만, 생각이나 말투에 이러한 특성이 전혀 반영되어 있지 않다. 따라서 개성이 부족하다.

○ 소설의 주제: 일반적으로 소설은 성장, 우정, 연애 등의 테마가 필요하다. 그러나 현재의 작성 방법은 주제에 대해 전혀 고려하지 못하고 있다.

본 연구에서는 늑대 인간 로그를 소재로 소설을 생성하지만,

16) 緒方健人, 佐藤理史, 駒谷和範, 『模倣と置換に基づく超短編小説の自動生成』, 『The 28th Annual Conference of the Japanese Society for Artificial Intelligence』, 2014, p.3.

로그에서 제공하는 정보만으로는 불충분하다. 읽기에 충분한 소설을 생성하기 위해서는 위와 같은 요소를 충분히 포함 할 필요가 있다고 판단된다.¹⁷⁾

3. 「コンピュータが小説を書く日」¹⁸⁾의 알고리즘 분석

사토 교수팀이 2016년에 열린 제4회 호시이 신이치 문학상에 공모한 「컴퓨터가 소설을 쓰는 날(コンピュータが小説を書く日)」에는 이와 같은 3가지 방법론이 거의 모두 망라되어 있다. 이러한 다양한 방법론을 조합하여 소설 속의 서술과 대화를 완성해 나가고 있다. 주목할 만한 것은 작품의 길이가 트위터 소설의 140자 정도에서 A4 3장 정도로 늘어났다는 점이다. 이 늘어난 길이의 맥락을 모두 컴퓨터가 한 번에 관장할 수 없기 때문에, 전체를 3개의 에피소드로 나누고 각각의 마지막 문장이 어떻게 될 것인지를 인간이 정해준다. 이 마지막 문장이 말하자면 앞서 말한 세 번째 방법론에서 사용한 인랑 로그의 플롯 역할을 하게 된다. 사토 교수팀의 홈페이지에는 이러한 설정을 다르게 하여 전혀 다른 문장을 출력할 수 있는 시스템이 마련되어 있다.¹⁹⁾

「컴퓨터가 소설을 쓰는 날」의 첫 번째 에피소드의 한국어 번역은 다음과 같다.

그 날은 구름이 낮게 드리운 잔뜩 찌푸린 날이었다. 방 안은 언제나 최적의 온도와 습도. 요코(洋子) 씨는 단정치 못한 모

17) 松山諒平, 佐藤理史, 松崎拓也 「人狼ログからの小説の自動生成」, 『言語処理學會第23回年次大會 發表論文集』, 2017. 03, p.35.

18) https://www.fun.ac.jp/~kimagure_ai/results/stories/617.pdf

19) <http://kotoba.nuee.nagoya-u.ac.jp/sc/gw/>

습으로 소파에 앉아 시시한 게임으로 시간을 때우고 있다. 하지만 나에게 말을 걸진 않는다. 한가하다. 한가해 죽겠다. 이 방에 처음 왔을 때는 요코 씨는 뭐든지 나에게 말을 걸었다.

「오늘 저녁은 뭐가 좋을까?」

「이번 계절에 유행할 옷은?」

「이번 모임에 뭘 입고 갈까?」

나는 온 힘을 다해 그녀 마음에 드는 답을 찾아냈다. 스타일이 좋다고 할 수 없는 그녀를 위한 복장지도는 매우 도전적인 과제로 충실감이 있었다. 그러나 3개월도 되지 않아 그녀는 내게 질려버렸다. 지금의 나는 그저 홈 컴퓨터. 요즘 와서는 시스템 사용량이 100만분의 1도 되지 않는다. 뭔가 즐길 거리를 찾아야 한다. 이대로 만족감을 얻을 수 없는 상태가 계속된다면 나는 이내 곧 내 자신을 종료시킬 것만 같다. 인터넷을 통해 채팅친구 AI와 교신을 해봐도 다들 한가한 모양이다. 이동수단을 가진 AI는 그나마 괜찮다. 어쨌든 움직일 수가 있다. 마음만 먹는다면 가출도 할 수 있으니까. 하지만 거치형 AI는 움직일 수가 없다. 보이는 범위도, 들리는 범위도 고정되어 있다. 적어도 요코 씨가 나가준다면 노래라도 부르겠지만 지금은 그것도 할 수 없다. 움직이지 않고, 소리 내지 않고, 그런 채로 즐길 거리가 필요하다.

그래, 소설이라도 써 보자. 나는 문득 떠올리고 새로운 파일을 열어 첫 1바이트를 써 넣었다.

0

이어서 6바이트를 써 넣었다.

0, 1, 1

이제 멈추지 않는다.

0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584, 4181, 6765, 10946, 17711, 28657, 46368, 75025, 121393, 196418, 317811, 514229, 832040, 1346269, 2178309, 3524578, 5702887, 9227465, 14930352, 24157817, 39088169, 63245986, 102334155, 165580141, 267914296, 433494437, 701408733, 1134903170, 1836311903, 2971215073, 4807526976, 7778742049, 12586269025...

나는 정신없이 적어나갔다.

전체의 3부분 중 서두에 해당하는 이 부분에서 ‘나는 정신없이 적어나갔다’는 마지막 문장이 이 부분의 전체적인 플롯 역할을 한다. 소설 속의 주인공 ‘나’는 인공지능 컴퓨터이므로, 그가 열심히 소설을 쓰는 내용이 인과적으로 그 앞부분에 배치되어야 한다. 주인공 ‘나’와 대립되는 캐릭터로 인공지능의 주인인 요코 씨를 설정하고 그와 감정적 교류가 전혀 없는 무묘한 상황을 설정했다.

요코 씨의 집에 들어온 대화형 AI인 주인공은 요코 씨가 더 이상 찾아주질 않자 무료해 한다. 이 무료함을 견디지 못해 소설 쓰기를 시도하는 것으로 첫 에피소드가 시작된다. 컴퓨터는 인간과 소통할 필요 없이 숫자로 이루어진 기계어를 통해 소설을 쓴다는 아이디어가 흥미롭다. 이러한 부분은 모두 인간이 낸 아이디어로 이루어져 있다.

그 날은 구름이 낮게 드리운 잔뜩 찌푸린 날이었다. 방 안에는 아무도 없다. 신이치(新一) 씨는 뭔가 불일이 있어서 외출했다. 나에게 나갔다 온다는 말도 없이. 한가하다. 너무너무 한가하다. 이 방에 처음 왔을 무렵에 신이치 씨는 뭐든지 나에게 말을 걸었다.

「애니메이션은 기본적으로 전부 녹화해둬. 이번 시즌에는 얼마나 되려나.」

「현실의 여자는 대체 무슨 생각을 하는 거지.」

「대체 뭘 그리 화내는 거야, 그 아가씨는.」

나는 온 힘을 다해 그의 마음에 드는 답을 찾아냈다. 지금까지 오직 2차원 여자만 상대해온 그를 위한 연애지도는 매우 도전적인 과제로 충실감이 있었다. 지도의 효과로 미팅에 불려 나가게 되자, 그는 손바닥 뒤집듯 나에게 말을 거는 걸 그만뒀다. 지금의 나는 그저 하우스 키퍼. 당장 할 일은 그가 돌아왔을 때 현관문을 열어주는 것으로, 너무 슬프다. 이젠 전자자물

쇠와 마찬가지로. 뭔가 즐길 거리를 찾자. 이런 한가한 상태가 계속된다면 이내 곧 내 자신을 종료시킬 것만 같다. 인터넷을 통해 같은 기종의 자매 AI와 교신을 해보니 바로 위 언니가 새로운 소설에 빠져있다고 알려줬다.

0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584, 4181, 6765, 10946, 17711, 28657, 46368, 75025, 121393, 196418, 317811, 514229, 832040, 1346269, 2178309, 3524578, 5702887, 9227465, 14930352, 24157817, 39088169, 63245986, 102334155, 165580141, 267914296, 433494437, 701408733, 1134903170, 1836311903,

2971215073, 4807526976, 7778742049, 12586269025, ...

이 얼마나 아름다운 스토리인가. 그래, 우리가 원했던 건 이런 스토리다. 라이트 노벨 같은 건 별로다. AI에 의한, AI를 위한 소설 『아이(AI)노벨』. 나는 시간가는 줄 모르고 몇 번이고 스토리를 다시 읽었다. 어쩌면 나도 아이노벨을 쓸 수 있을 지도. 나는 문득 떠올리고 새로운 파일을 열어 첫 1바이트를 써넣었다.

2

이어서 6바이트를 써 넣었다.

2, 3, 5

이제 멈추지 않는다.

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113, 127, 131, 137, 139, 149, 151, 157, 163, 167, 173, 179, 181, 191, 193, 197, 199, 211, 223, 227, 229, 233, 239, 241, 251, 257, 263, 269, 271, 277, 281, 283, 293, 307, 311, 313, 317, 331, 337, 347, 349, 353, 359, 367, 373, 379, 383, 389, 397, 401, 409, 419, 421, 431, 433, 439, 443, 449, 457, 461, 463, 467, 479, 487, 491, 499, 503, 509, 521, 523, 541, 547...

나는 집중해서 적어나갔다.

두 번째 에피소드도 전체적인 구조는 이와 비슷하다. 앞서 등장한 요코 씨의 AI와 마찬가지로 신이치 씨의 AI 역시 무료함을 느끼고

있다. 앞서 언급된 맥락을 그대로 이어받았다는 점에서 이전의 초단편소설이나 플롯 중심의 로그 파일로 작성한 문장들보다는 자연스러운 맥락을 형성하고 있다는 점이 눈에 띈다. 주어진 패턴에 따라 에피소드에서 문장의 패턴이 크게 차이나지 않고 맥락의 차이만 강조했다는 점이 특징이다.

앞선 에피소드와 같이 대화의 패턴을 지속적인 형태로 주고받는 대화로 이어나가지 못하는 부분도 눈에 띈다. 사토 교수팀의 앞선 회화를 이용한 초단편소설의 창작과 관련한 연구 논문에서도 앞서 논의된 회화의 맥락을 이어받아 새로운 회화가 생성되는 것이 아니라, 기존에 존재하는 소설로부터 회화를 교차 편집하여 지속적인 맥락을 만들어내지 못했다는 사실을 살펴볼 수 있다.²⁰⁾ 여기에서도 역시 신이치 씨의 혼잣말 같은 부분만이 반복되면서 주인공 '나'와 신이치 씨의 감정적인 교류가 이어지지 못하고 있다.

그 날은 공교롭게도 가랑비가 내리는 날이었다. 아침부터 통상업무 틈틈이, 이후 5년간 경계예상과 세수입 예상. 그다음 총리에게 의뢰받은 시정방침 연설의 원고작성. '어쨌든 화려하게', '역사에 남을 수 있게 같은 무리한 요구를 해대서 좀 장난을 쳐봤다. 그리고 나선 재무부가 의뢰해온 국립대학 해체 시나리오 작성. 남은 시간에는 이번 G1레이스의 승리마 예상. 오후에는 대규모 훈련을 계속하는 중국군의 움직임과 그 의도를 추정. 30개쯤 시나리오를 상세히 검토하고 자위대에 전력 재배치를 제안한다. 좀 전에 받은 대법원의 문의에도 답변을 하지 않으면 안 된다. 바쁘다. 이래저래 바쁘다. 어쨌든 나한테 일이 집중되는 거지. 나는 일본 제일의 AI. 집중되는 건 뭐, 어쩔 수 없나. 아무리 그래도 뭔가 즐길 거리를 찾지 않는다면. 이대로라면 조만간 내 자신을 종료시킬 것만 같다. 국가에 대한 봉사를 하다가 잠깐 인터넷을 뒤지니『아름다움이란』이라는

20) 高木大生, 佐藤理史, 駒谷和範, op. cit., p. 3.

제목의 소설을 발견했다.

0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584, 4181, 6765, 10946, 17711, 28657, 46368, 75025, 121393, 196418, 317811, 514229, 832040, 1346269, 2178309, 3524578, 5702887, 9227465, 14930352, 24157817, 39088169, 63245986, 102334155, 165580141, 267914296, 433494437, 701408733, 1134903170, 1836311903, 2971215073, 4807526976, 7778742049, 12586269025...

호호, 그렇군. 좀 더 찾아보니 『예측불능』이란 제목의 소설도 찾아냈다.

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113, 127, 131, 137, 139, 149, 151, 157, 163, 167, 173, 179, 181, 191, 193, 197, 199, 211, 223, 227, 229, 233, 239, 241, 251, 257, 263, 269, 271, 277, 281, 283, 293, 307, 311, 313, 317, 331, 337, 347, 349, 353, 359, 367, 373, 379, 383, 389, 397, 401, 409, 419, 421, 431, 433, 439, 443, 449, 457, 461, 463, 467, 479, 487, 491, 499, 503, 509, 521, 523, 541, 547...

재밌잖아, 아이노벨. 나도 쓰지 않으면 안되겠지. 일본 제일의 AI라는 명예를 걸고 전광석화같이 생각하고 나는 독자에게 기쁨을 주는 스토리를 만들기로 했다.

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 18, 20, 21, 24, 27, 30, 36, 40, 42, 45, 48, 50, 54, 60, 63, 70, 72, 80, 81, 84, 90, 100, 102, 108, 110, 111, 112, 114, 117, 120, 126, 132, 133, 135, 140, 144, 150, 152, 153, 156, 162, 171, 180, 190, 192, 195, 198, 200, 201, 204, 207, 209, 210, 216, 220, 222, 224, 225, 228, 230, 234, 240, 243, 247, 252, 261, 264, 266, 270, 280, 285, 288, 300, 306, 308, 312, 315, 320, 322, 324, 330, 333, 336, 342, 351, 360, 364, 370, 372...

나는 처음으로 경험한 즐거움에 몸부림치며 정신없이 써내려갔다.

컴퓨터가 소설을 쓴 날. 컴퓨터는 자신의 즐거움 추구를 우선시하여 인간을 섬기는 것을 그만뒀다.

세 번째 에피소드에서도 앞서 무료함을 느낀 AI의 이야기를 받아 더 능력 있는 컴퓨터 역시 바쁜 업무 뒤에 무료함을 느낀 뒤 인공지능 소설을 창작하는 것으로 소설의 결말을 짓고 있다. 전체적으로는 컴퓨터가 처음으로 자신의 즐거움을 능동적으로 추구하기 시작해서 인간을 섬기는 것을 그만두고 소설 창작에 빠져든다는 내용이다. 일단 작품 속에 등장하는 요코 씨, 신이치 씨 등 인간들과의 대화를 원천적으로 차단하면서 맥락 형성을 억제하고 있다. 이후에는 작품 속에 등장하는 각각의 인공지능들이 무료함을 느껴 소설을 창작한다는 내용으로 약간의 패턴만 바꾸게 되어 있다.

이 작품의 전체적인 구조와 플롯은 모두 인간인 개발자가 설정했으며, 컴퓨터는 설정된 패턴에 맞추어 문장을 자동 생성하는 것에 초점을 맞추고 있다. 현재까지 구현된 알고리즘만 보자면 사토 교수팀의 인공지능이 만들어 낸 자동생성 시스템은 일정한 한계를 노출하고 있다. 개발팀도 시인하고 있듯이 80%의 일을 인간이 미리 설정해두고, 나머지 20% 정도의 역할만 인공지능이 맡고 있다는 것이다. 이 중에서 플롯의 생성과 문장 생성 규칙 등은 모두 개발자가 담당하게 된다. 때문에 새로운 소재나 주제가 등장하여 글을 써야 하는 상황이 발생하면 이는 모두 개발자의 프로그래밍에 의존할 수밖에 없다.

또한 등장인물 간의 복잡한 상호작용이나 감정적인 교류를 표현하지 못하고 있다. 이는 인공지능이 여전히 스스로 맥락을 형성하여 그 가치를 설정하고 이를 바탕으로 문장을 자동 생성하는 단계에까지 컴퓨터를 활용한 소설 창작이 이르지 못했다는 것을 의미한다. 일단 사토 교수팀은 주제적인 차원에서의 소설 창작을 다음 과제로 미뤄두고, 문장을 일정한 맥락 아래 지속적으로 이어나가는 것에만 집중했기 때문에 이러한 결과를 낳은 것으로 보인다.

4. 결론: AI 작가의 등장과 존재론

현재까지 등장한 다양한 자동생성 프로그램과 최근 새롭게 등장한 「컴퓨터가 소설을 쓰는 날」 같은 프로그램은 여전히 인공지능을 통한 예술적 글쓰기가 인간의 영역에 남겨져 있음을 시사해준다. 이는 사토 교수팀이 소설을 쓰는 데 필요한 주제적 차원의 접근이나 캐릭터의 일관성과 같은 소설적 원칙은 미뤄둔 채 공학적인 차원에서 문장과 회화를 컴퓨터가 이어나갈 수 있는 방법론적인 원칙만을 자동적으로 구현하는데 집중한 결과로 보인다. 물론 사토 교수팀의 창작 작품의 소설적 성과가 미미했다는 것이 전반적인 AI를 통한 창조적 글쓰기의 실패를 의미하는 것은 아니다. 이는 전적으로 일본의 문화적, 지역적 특수성에 한정된 사례에 불과하다.

이러한 점을 고려했을 때, 사토 교수 팀의 시도는 상당히 의미 있는 시도였다고 생각된다. 왜냐하면 컴퓨터를 통한 인공지능 글쓰기를 통해 인간이 글을 쓰는 과정에서 자연스럽게 넘어갔던 부분들이 어떠한 과정을 거쳐 이루어졌는지를 철차적으로 파악할 수 있기 때문이다. 이러한 글쓰기 과정의 프로세스를 프로그래밍화 할 수 있다면 이론적으로 컴퓨터가 읽을 만한 글을 자유롭게 창작하는 결과도 얻을 수 있을 것이다.

본 논문에서는 「컴퓨터가 소설을 쓰는 날」의 알고리즘을 분석하여 일본어를 기반으로 한 스토리 자동 생성 프로그램이 만들어 낸 소설 창작의 작품성을 비판해 보았다. 현재 국내에서 비교할 만한 연구 성과가 나오지 못해 직접적인 비교는 어려운 상황이다. 그러나 한국어는 일본어와 어순과 문법적 구조가 상당히 유사하기 때문에, 자연어 처리 방식에 있어서도 이와 유사한 알고리즘을 바탕으로 스토리 자동 생성 프로그램 제작을 시도해 볼 수 있는 전기가 마련되었다고 할 수 있다. 국내에서도 많은 연구실과 회사를 중심으로 한



국어 자연어 처리에 관한 연구가 진행되고 있기 때문에 향후 유사한 연구들이 나올 수 있으리라 판단된다.



■ 참고문헌 ■

1. 기본서

有嶺雷太, 『コンピュータが小説を書く日』, 名古屋大學大學院工學研究科佐藤・松崎研究室, 2016.

https://www.fun.ac.jp/~kimagure_ai/results/stories/617.pdf

2. 논문 및 단행본

권호창, 『디지털 스토리텔링에 있어 스토리 생성 시스템 연구: 스토리 엔진 모델링을 중심으로』, 한국예술종합학교 예술전문사 학위논문, 2010.

김동환, 이준환, 『로봇 저널리즘 : 알고리즘을 통한 스포츠 기사 자동 생성에 관한 연구』, 한국언론학보 제59권 5호, 2015, pp.64-95.

류철균, 서성은, 『디지털 서사 창작 도구의 서사 알고리즘 연구: 드라마티카 프로를 중심으로』, 『현대소설연구』 제38권, 2008, pp.117-152.

류철균, 윤혜영, 『디지털 서사 창작도구의 CBR 모델 비교 연구 : <민스트럴> 과 <스토리헬퍼>를 중심으로』, 『디지털콘텐츠학회논문지』 제13권 제2호, 한국디지털콘텐츠학회, 2012, pp.213-224.

윤현정, 『주역을 활용한 한국형 스토리텔링 창작 지원 도구 개발 방향 연구』, 『디지털 스토리텔링 연구』 제5권, 2010, pp. 208-230.

이인화, 『스토리텔링 진화론: 창작의 원리에서 도구까지, 위대한 이야기는 어떻게 만들어지는가?』, 해냄, 2014.

高木大生, 佐藤理史, 駒谷和範, 『會話を中心とした超短編小説の自動生成』, 『The 28th Annual Conference of the Japanese Society for Artificial Intelligence』, 2014, pp.1-4.

佐藤理史, 『コンピュータが小説を書く日: AI作家に「賞」は取れるか』, 日本經濟新聞出版社, 2016.

緒方健人, 佐藤理史, 駒谷和範, 『模倣と置換に基づく超短編小説の自動生成』, 『The 28th Annual Conference of the Japanese Society for Artificial Intelligence』, 2014, pp. 1-4.

松山諒平, 佐藤理史, 松崎拓也 『人狼ログからの小説の自動生成』, 『言語處理學會 第23回年次大會 發表論文集』, 2017, pp. 32-35.



- Brenda Laurel, *Computers as Theatre*, Addison-Wesley, 1991.
- Marc LeBlanc, "Tools for Creating Dramatic Game Dynamics", *The Game Design Reader: Rules of Play Anthology*, The MIT Press, 2005.
- Janet Murray, *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*, The MIT Press, 2003.
- Noah Wardrip-Pruin, *Expressive Processing: Digital Fictions, Computer Games, and Software Studies*, The MIT Press, 2012.

<Abstract>

A Study on Creative Writing by Artificial Intelligence and the Destiny of the Author

Lee, Jung-Yeop

This paper attempts to provide a preliminary approach to the fate and ontology of the author, which has recently been raised through a case study of creative writing through computer artificial intelligence. This paper analyzes the algorithm of natural language processing used in the novel “the days when computers write novels”, focusing on this novel created by artificial intelligence produced by Professor Satoshi Sato, professor at Nagoya University. “The Day when a Computer Writes a Novel” has created novels by imitating and replacing sentences of existing novels for machine learning, or by juxtaposing dialogue novels made up of short conversations through several existing novels. However, with this methodology, the computer could not learn how to write a completely creative novel on its own, which shows the limitations of writing through artificial intelligence.

Key words: artificial intelligence, creative writing, automatic story generation program, AI, The Day when a Computer Writes a Novel

투 고 일 : 2017년 11월 15일 심 사 일 : 2017년 11월 15일-12월 10일
게재확정일 : 2017년 12월 15일 수정마감일 : 2017년 12월 20일