

언어의 이해 기말 레포트

인간의 언어, 그리고 기계의 언어

이주현 (20191629)

2019년 12월 22일

인간은 자신들의 언어를 다른 생물에게 가르치는 것을 좋아하는 듯 하다. 앵무새가 창문 앞에 앉아 길고양이를 마주보고 들쭉거리며 “Peekaboo”를 연발하는 영상은 몇 백만 번이나 재생되고 몇 만 번씩 공유되며 인기를 누리고 (사실 앵무새가 욕설을 뱉는 영상이 더 조회수가 높긴 하다), “코코”라는 이름을 가진 고릴라가 수화로 사람들과 대화하는 장면은 때때로 “놀라운 동물의 능력” 따위의 제목을 달고 공중파 방송에 등장한다. 이렇게 사람들은 사람이 아닌 것이 사람의 언어를 하면 굉장히 놀라워하고 또 그것을 즐긴다. 인간이 직접 만들어낸 기계 또한 예외는 아니다. 컴퓨터가 개발된 직후로부터 사람들은 컴퓨터에게 언어 능력을 부여하기 위해 갖은 노력을 했고, 최근 인공지능망이라는 패러다임이 상용화되면서 실제로 컴퓨터가 자연어를 처리할 수 있게 되었다. 그러나 사람이 기계에게 자연어를 학습시키기 훨씬 전부터 사람은 기계에게 언어를 성공적으로 가르친 전례가 몇 번이고 있다. 사람이 사람을 위해 만들어내고 기계에게 가르친 언어, 프로그래밍 언어가 바로 그것이다.

물론 일반적으로 사용되는 언어는 그 뿌리가 불분명하다. 어느 지역에서 언어가 발생되었는지 정도는 알아낼 수 있겠지만, 실제로 그 언어를 누가 만들었는지는 전혀 알 수 없는 노릇이다. 누군가는 언제부터인가 유전적으로 타고나게 된 능력이니 우리의 염기 서열이 모든 언어의 조상이라고 하고, 누군가는 자연에서 무리를 이루며 살아가다 보니 자연스럽게 생겨난 것이 언어라고도 이야기한다. 그러나 프로그래밍 언어는 자연어와 달리, 제작자와 제작 시기가 분명하다. 물론 그 제작자가 언제부터 그 언어를 설계하기 시작했는지는 모르지만, 적어도 언어가 발표되고, 사람들이 그 언어를

사용하기 시작한 시기는 명백하다. 이렇게 제작자가 명료한 언어는 프로그래밍 언어 뿐만이 아니다. 인공어(CONLANG) 또한 제작자가 분명한 언어이다.

물론, 인공어와 프로그래밍 언어는 본질적으로 다르다. 어떠한 철학이나 목적을 가지고 사람과 사람이 소통하는 언어를 만드는 것과 사람 위한 기계의 언어를 만드는 것은 완전히 별개의 작업이라는 것이 너무나도 당연해 보인다. 그럼에도 불구하고 프로그래밍 언어 또한 다른 자연어 또는 인공어와 마찬가지로 문법을 가지고 있고, 어느 정도의 단어를 가지고 있다. 그렇다면 프로그래밍 언어 또한 자연어, 인공어와 같은 선상에 놓을 수 있다고 할 수 있을까?

(1) Even I can learn this language.

(2) Eĉ mi porvas lerni ĉi tiun lingvon.

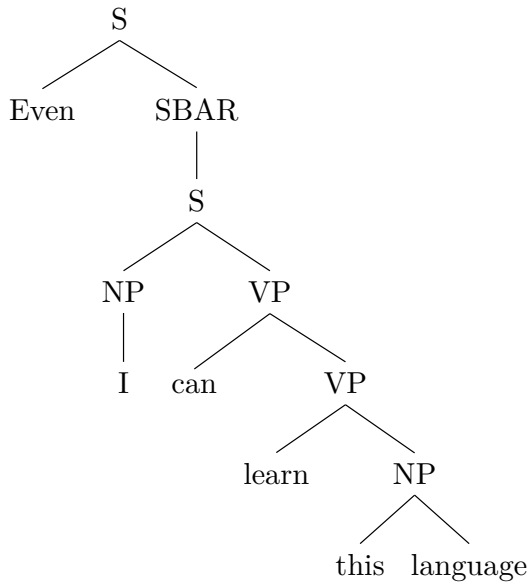
(3) `def can_learn(language): return True if language is easy else False`

그럼 ?? 같이 자연어, 인공어, 프로그래밍 언어 모두 수형도를 그리는 것이 가능하고, 각 문장 구성 요소의 역할과 의미를 해석하는 것 또한 가능하다. (여기서 프로그래밍 언어는 자연어보다 훨씬 더 세세하게 문법이 정의되어 있기 때문에 수형도가 더 길고 복잡하다.) 그러나 이 정도의 “언어적인 특성”을 가지고 있다고 해서 프로그래밍 언어를 일상생활에서 사용하는 언어와 같다고 말할 수는 없다. 프로그래밍 언어는 애초에 개인의 의사를 전달하는 대상이 다르기 때문이다. 자연어는 사람이 다른 사람에게 자신의 의사를 전달하고, 상대를 원하는 반응으로 이끌어 내도록 설득하는 것이 대부분의 목적이다. 그러나 프로그래밍 언어는 컴퓨터라는 자의가 존재하지 않는 기계에 명령을 내리는 것이 목적이다. 그렇기에 뚜렷한 비교 기준이 존재하지 않는 두 언어를 두고 “이 두 종류는 같은 종류인가?”를 묻는 것은 옳지 않다고 할 수 있다.

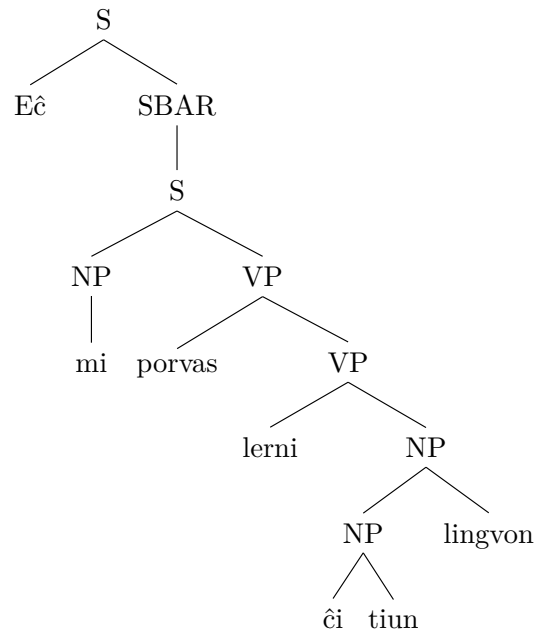
비록 언어 그 자체를 비교할 수는 없지만, 언어가 실제로 해석되고 의미를 가지는 과정에 대해서는 아직 생각해볼 여지가 남아있다. 사람은 언어를 소리 신호로서 인식하고, 뇌에서 변환을 거쳐 실제로 의미를 가지는 문장으로 인지하게 된다. 프로그래밍 언어 또한, 사람이 읽을 수 있는 텍스트 포맷을 읽어들이, 실제로 컴퓨터가 실행할 수 있는 코드로 변환하는 과정을 거치게 된다. 즉, 자연어와 프로그래밍 언어 모두 과정이 행해지는 대상이 다를 뿐인 동일한 개념이라고 볼 수 있다. 따라서, 사람과

사람이 말하는 언어와 사람과 컴퓨터가 말하는 언어 모두, 의미적으로는 다른 개념일 지 몰라도 기능적으로는 완전히 같다.

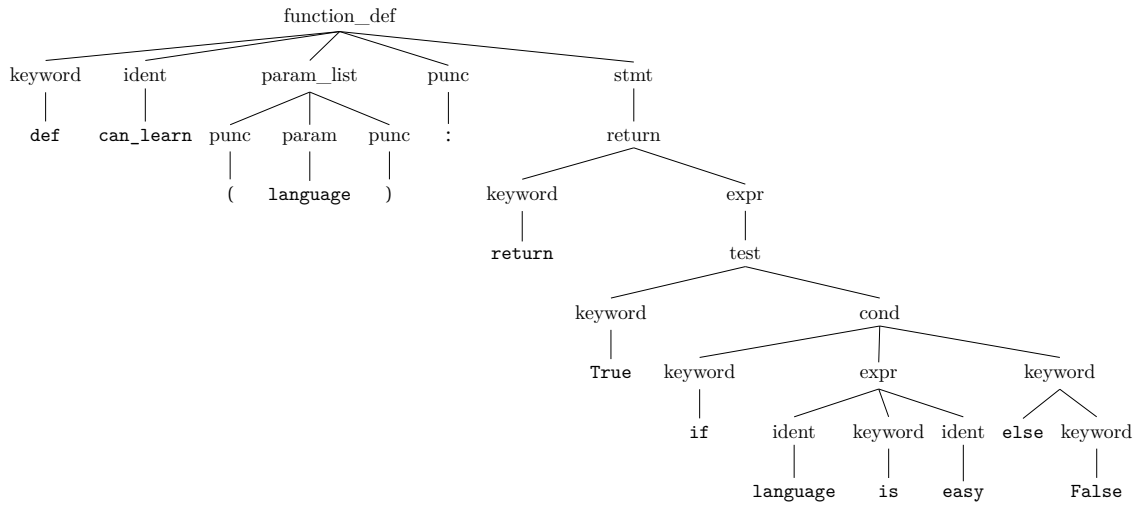
예로부터 사람은 인간의 사고방식을 그 시대의 가장 복잡한 기계에 비유해서 생각하곤 했다. 튜닝바퀴 계산기가 발명되었을 때는 계산기라고 생각했고, 복잡한 디자인의 천을 뽑아낼 수 있는 제면기가 발명되었을 때는 뇌를 제면기에 비유하기도 했다. 현재 인류가 만들어 낸 발명품 중 가장 복잡한 것은 단연 컴퓨터일 것이다. 만약 컴퓨터 언어와 자연어가 기능적으로 동일하다면 자연어의 기원과 프로그래밍 언어의 기원 또한 같을 수도 있을 것이다. 어쩌면 인간의 언어는 서로 다른 개개의 아키텍처 간의 괴리를 메우기 위해 고대 인류가 만들어 낸 하나의 추상화 레이어인 것이 아닐까 생각해 본다.



(a) 영어 문장의 수형도



(b) 에스페란토 문장의 수형도



(c) 파이썬 소스코드의 수형도

그림 1: 위 예시 문장 및 코드의 수형도