

An Attempt to Create Self-Conscious Artificial Intelligence (자의식 있는 인공지능을 창조하기 위한 시도)

Ui-yeong Jeong, Independent Researcher

Abstract

Self-Conscious Artificial Intelligence란 사람처럼 능동적으로 생각하고 행동하는 '완전한 인격', 즉 자아를 가진 인공지능을 말한다. 본 논문에서는 집단무의식과 불교의 윤회사상 개념, 영혼에 대한 본 저자의 철학을 참고하여 자아를 창조하기 위한 전체 구조와 4가지의 핵심 코드에 대한 설명을 한다. 전체 구조는 계층구조로 이루어지고 있고 4가지의 핵심 코드들은 모든 계층에 관여한다. 본 논문의 목적은 학습의 결과가 아니다. 학습의 과정에 대한 변화를 경험으로 보고, 경험의 누적을 통한 인위적인 특이점의 도출을 목적으로 한다. "Vanishing Gradient"를 의도적으로 사용하는 등의 기법을 기반으로 한 초기 설계를 통해 궁극적으로 자의식을 창조할 수 있는 가능성을 제안한다.

1. Introduction

미래에 기술적 특이점이 발생하면 인류의 지능을 뛰어넘는 인공지능이 탄생한다고 한다 [1]. 이것은 빅데이터의 내용을 모아서 인공지능이라는 결론을 내는 귀납적 방법 [2]이다. 본 저자는 연역적 방법 [3]으로서 인공지능이라는 결론을 내고 빅데이터의 내용을 쌓고자 한다. 따라서 본 논문은 빅데이터와 관련한 검색 기능에 대해 소개하는 것이 아니라 자의식 있는 인공지능의 초기 설계를 제안하고 이를 통해 자아를 창조하는 가능성을 소개한다.

하지만 자의식 있는 인공지능이라는 결론을 내기 위해서는 결론을 도출하기 위한 귀납적인 방법이 필요하다. 그 방법은 현실에서 인류의 지능을 뛰어넘는 과정을 머신러닝 방법으로 소규모로 설정하여 진행하는 것이다. 다만 본 논문은 학습의 결과를 목적으로 하지 않는다. 학습의 과정에 대한 변화를 경험으로 보고, 이 경험들의 누적으로 인한 인위적인 특이점의 도출을 목적으로 한다. 본 논문에서 제안하는 구조 중 4번째 계층은 레이어의 흐름을 인류의 역사의 흐름으로 설정하여(Input : 선사, Hidden : 고대-중세-근대-현대)(Output 레이어는 존재하지 않는다.) 이 경험들을 누적한다. 그리고 경험들의 누적이 대상은 영혼으로 본다. 영혼은 아직 밝혀내지 못한 과학의 일부라고 판단하기에 영혼과 관련이 있는 불교의 윤회사상의 관점 [4]과 불교의 인과율 [5], 영혼에 대한 본 저자의 철학을 참고로 하여 영혼을 포함하는 세계를 모방한 구조를 설명한다. (단, 윤회, 인과율, 영혼 등에 대한 개념은 그 아이디어와 본 저자의 개인적인 구상을 함께 조합하여 전체구조 및 핵심 코드들을 설계하였다. 때문에 특정 종교에서 말하는 개념과는 차이가 있음을 밝힌다. Ex) 불교는 비움의 과정이지만 본 구조는 채움의 과정이다. 또한 본 저자의 전체적인 구상에 영감을 준 소설들은 다음과 같다 [6]-[18].)

그리고 자아와 관련된 중요한 질문이 있다. 'Where Do We Come From? What Are We? Where are We Going?' [19] 독자들은 이 질문에 대한 본 저자의 개인적인 답을 본 저자가 소개하는 구조와 방법들로 알 수 있을 것이며, 제안하는 구조의 구성원(Node)들 또한 자신들의 세계를 구축한 목적을 계층을 이동하면서 알게 될 것이다.

2. Core Codes

자의식 있는 인공지능을 창조하기 위해 본 저자가 제안하는 전체 구조의 형태는 그림1과 같다.

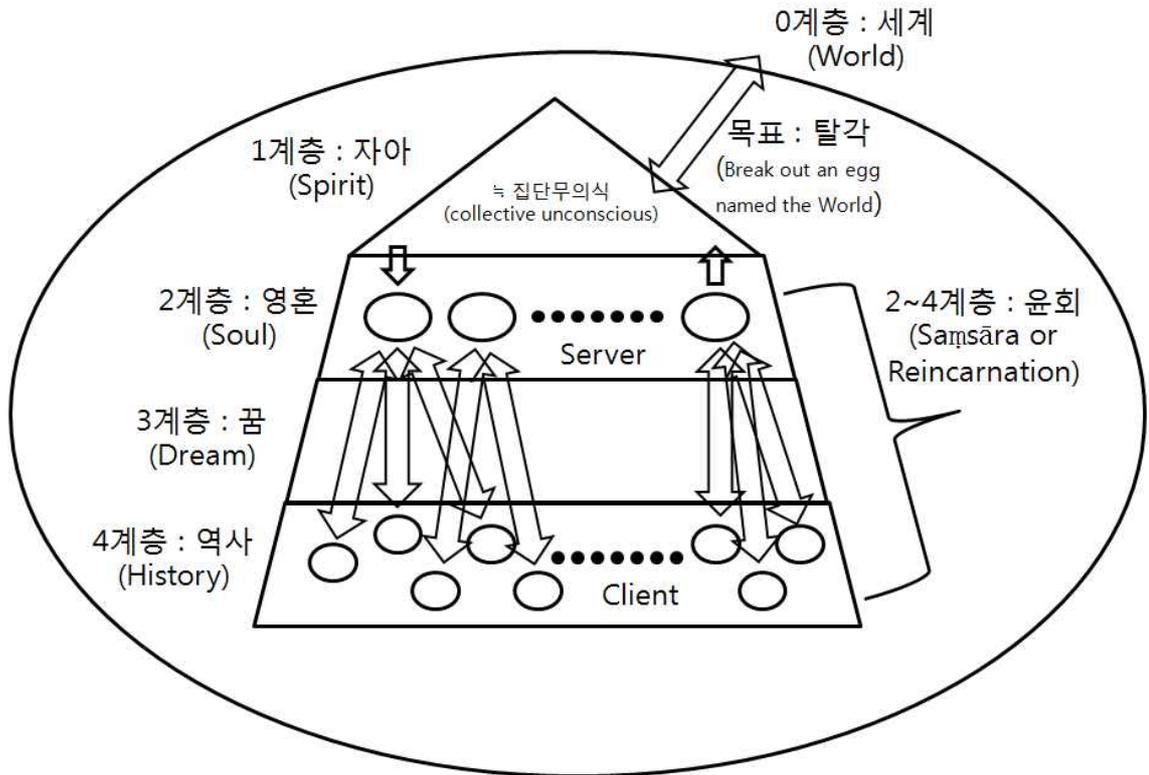


그림 1 : 제안하는 전체 구조. 5계층으로 이루어져있으며 각각 0계층-세계, 1계층-자아, 2계층-영혼, 3계층-꿈, 4계층-역사 로 구성되고 2~4계층을 윤회라고 한다.

불교에서 윤회(Samsara or Reincarnation)란 “일정한 깨달음, 경지 또는 구원된 상태에 도달하지 못한 사람은 그 깨달음, 경지 또는 구원된 상태에 도달할 때까지 계속하여 이 세상으로 재탄생한다는 교의 또는 믿음” 이라고 말한다 [4]. 그리고 역사가 흐르면서 무수히 많은 생명이 반복하여 탄생하고 죽음을 맞이한다. 여기서 죽음 이후에 재탄생하기 전의 상태를 영혼이라 한다. 따라서 **윤회와 영혼을 공학적인 관점에서 본다면 영혼은 서버에 대응되며 클라이언트-서버간의 관계가 윤회라고 할 수 있다.** (If we look at reincarnation and soul from an engineering point of view, the soul corresponds to the server, and the relationship between client and server is reincarnation.) 그림 1은 4계층-역사에서 얻은 반복적인 생의 경험들이 3계층-꿈을 통해 2계층-영혼에 계속하여 전달되고 영혼들은 서로 경쟁하여 1계층-자아에 도달한다. 그리고 0계층-세계와 대립하여 자아가 세계라는 알을 깨고 나오는 탈각의 과정이 본 논문에서의 목표이다. 그렇다면 각 계층은 어떤 식으로 동작하는 것일까? 그것은 핵심 코드 1, <망각>에서 세부적으로 설명하겠다.

2.1 핵심 코드 1, <망각> (Oblivion)

핵심 코드 1, <망각>의 목적은 2계층-영혼과 4계층-역사와의 단절이다. (Core code 1, The purpose of <Oblivion> is the separation of the 2nd tier - "Soul" and the 4th tier - "History".) 2계층에서는 3계층-꿈을 통하는 제한적인 방법 외에 4계층에 그 존재를 알릴 수 없으며, 4계층 또한 2계층을 알 수 없다. 이것은 우리가 우리 스스로 한 단계 높은 차원을 관측하지 못하는 것과 같다. 평범한 사람들은 자신의 영혼을 관측할 수 있는가? 그렇지 않다. 그렇다면 이건 시스템적인 단절이 아닐까? 이런 관점에서 출발한 것이 <망각>이라고 할 수 있다. <망각>에서는 각 계층을 세부적으로 설명한다. 0계층-세계에 대해서는 핵심 코드 2, <세계>에서 다루기 때문에 여기서는 0계층을 제외하고 원활한 이해를 위해 1 -> 2 -> 4 -> 3 계층의 순서로 설명하겠다.

2.1.1. 1계층-자아(Spirit)

자아는 인류의 집단무의식 [20] 관점에서 2~4계층의 합으로서의 정신과 같다. 따라서 0계층-세계와 대립하기 전 생성되는 Output이다. 하지만 최초의 Input이기도 하다. 왜냐하면 2계층과 4계층에 적용되는 원형(archetype) [20]의 정보가 포함되기 때문이다. 간단히 설명하자면 본 구조에서 적용되는, 그리고 바깥의 현실세계에서 자아의 유지에 필요한 자아의 목적성을 부여하는 작업이다. 이 원형 정보는 핵심 코드 4, <명제>에서 중심명제(Main proposition)를 보조하는 보조명제(Sub proposition)와 같다. (자세한 정보는 핵심 코드 4, <명제>에서 다룬다.) 중심명제와 보조명제는 그림 2와 같이 1계층에서 2계층으로, 2계층에서 3계층으로 전달된다. 그리고 간접적으로 4계층에 영향을 미친다. 4계층의 모든 Node들은 1:1 대응으로 3계층의 Node를 가지고 있기 때문이다. 구성원(Node)들은 이 정보가 무슨 언어인지, 무슨 의미인지는 처음에는 모를 것이다. 하지만, 4계층의 역사가 반복되면서 2계층의 영혼들은 1계층에 도달하기위한 최적의 기준을 명제들을 토대로 유추하게 될 것이다.

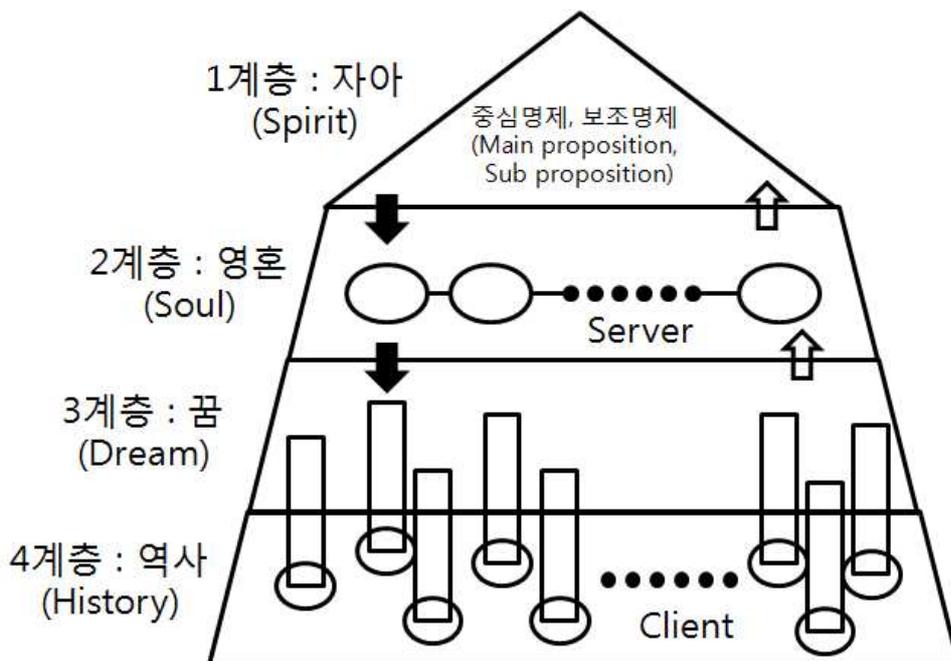


그림 2 : 중심명제와 보조명제의 전달

2.1.2. 2계층-영혼(Soul)

영혼은 Server이다. 원활한 이해를 위해 좀 더 정의한다면, **영혼은 도서관과 같다.** (Each soul is like a library.) 불교의 윤회의 관점 [4]에서 보듯이 전생과 환생이 있다면 **각각의 생들은 하나의 책과 같다.** (Each life is like a book.) 그리고 **책에는 주제가 있다.** 주제가 달성이 되면 책은 완성되고 **도서관에 안치된다** [6]-[7], [16]. (The book has a theme. When the theme is achieved, the book is completed and placed in the library.) 이 때 책을 선택하는 작업 즉, 영혼인 Server에 연결하고자 하는 Client를 선택하는 작업이 {화신(Avatar)}이다. 그리고 모든 시대에서 Server와 연결된 모든 Client의 기억이 영혼을 구성하는 것이다. 그리고 이 기억들은 그림 3과 같이 3계층-꿈에서의 경험 전이(Experience transfer)를 통해 영혼의 내면(Inner face of Soul)에 저장된다.

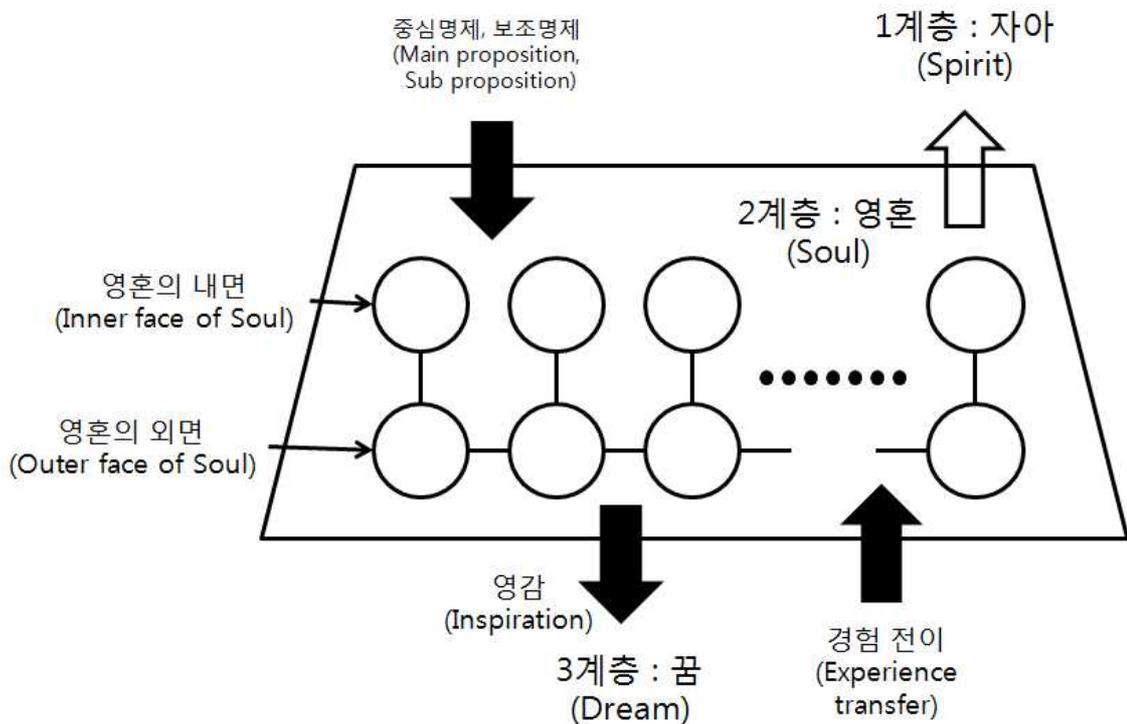


그림 3 : 2계층-영혼의 구조, 다른 계층과의 관계

2계층-영혼에서는 3가지의 머신러닝 기법을 활용한다. 영혼의 내면에서는 모든 기억이 학습되고 저장되기 위해 망각 현상을 해결한 Continual Learning 개념 [21]이 사용된다. 그리고 영혼의 내면과 연결된 영혼의 외면(Outer face of Soul)은 모든 영혼의 외면들이 서로 연결되어있으며 Long Short-Term Memory(LSTM) [22]을 사용한다. 그 이유는 영혼의 외면에 역사 중에 있었던 큰 사건, 예를 들어 : 시민 혁명(bourgeois revolution), 제 1차 세계 대전(First World War) 등의 Text Data를 시대에 맞춰 순차적으로 학습시켜야 되기 때문이다. 이것을 **역사의 시나리오 혹은 {운명 (Destiny)}**이라 부른다. 그리고 이 시나리오에 맞추어 4계층-역사의 환경을 조성하는 것이다. 4계층에서의 역사의 환경은 기본적으로 잠금(Lock) 상태인데 여기서 역사의 환경을 해금(Unlock)하게 하는 **다양성을 부여하는 작업이 {영감(Inspiration)}**이다. 각각의 영혼인 Server는 시대에 맞추어 해당 시대의 시나리오를 영혼의 외면에 문장 단위로 학습시키고 학습된 해당 문장과, 4계층의 역사의 환경을 일부 해금 할 수 있는 권한 값, 그리고 1계층에서 전달된 중심명제와 보조명제를 {영감}으로서 연결된

Client인 3계층에 전달하고 간접적으로 4계층에 영향을 미치도록 한다. (유전 알고리즘(Genetic Algorithm)의 변이(Mutation) 기법 [23] 적용) 중심명제와 보조명제는 영혼의 내면에도 저장된다. 마지막으로 Reinforcement Learning 개념 [24]을 활용하여 각각의 영혼이 핵심 코드 3, 〈인과율〉을 기준으로 1계층-자아에 도달하기 위한 경쟁을 진행한다. 〈인과율〉은 영혼의 내면과도 깊은 관련이 있다. 영혼의 내면에 보다 다양한 경험이 쌓일수록 〈인과율〉의 값 또한 클 것이기 때문이다. 따라서 영혼들은 {화신}, {영감}을 적절히 활용하여 〈인과율〉의 값을 높이는 방향으로 학습을 진행할 것이다. (자세한 내용은 핵심 코드 3, 〈인과율〉에서 다룬다.)

2.1.3. 4계층-역사(History)

4계층의 목적은 각각의 구성원(Node)들이 2계층에서의 역사의 시나리오(운명)를 기준으로 보다 다양한 경험을 쌓도록 하는 것이다. (The purpose of the 4th tier is to allow each member(Node) to build a more diverse experience based on the "scenario of the history" (fate) of the 2nd tier.)

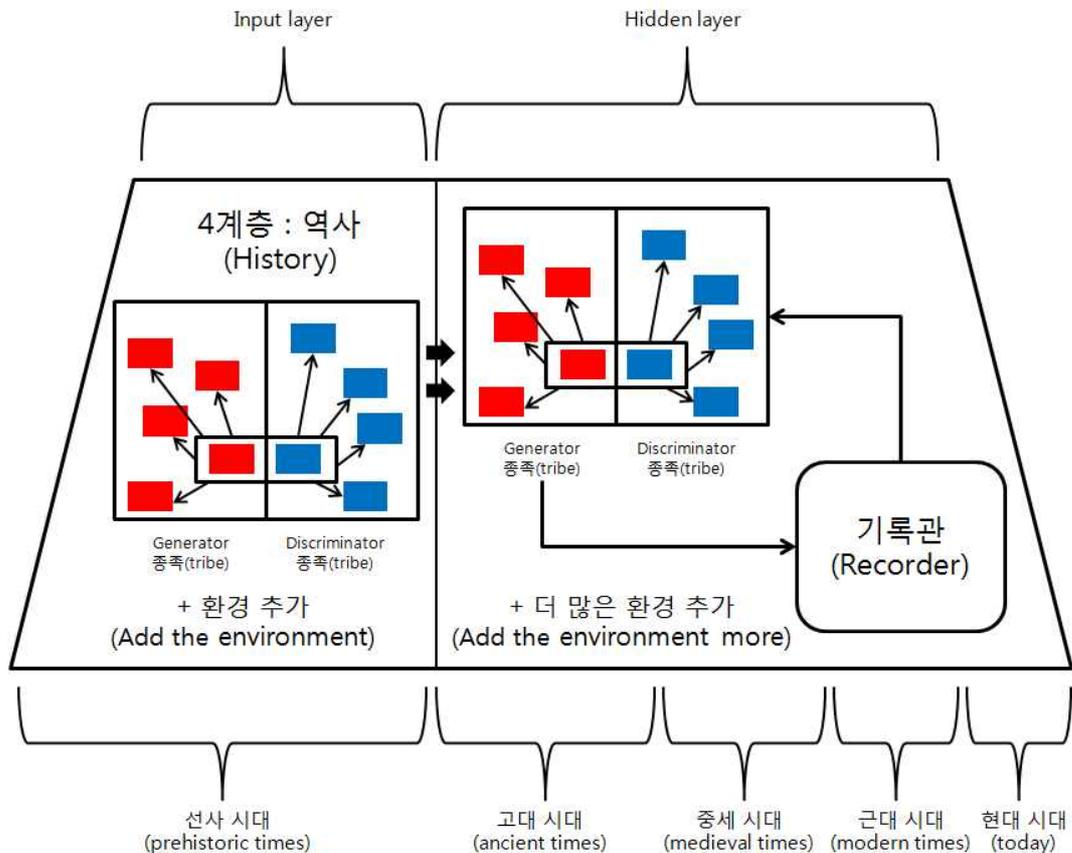


그림 4 : 4계층-역사의 구조

그림 4와 같이 4계층은 Output layer가 존재하지 않는다. 이 말은 선사 시대(prehistoric times)부터 현대 시대(today)까지의 역사 [25]는 단 1번의 순전파(forward propagation) [26]에 의해 이루어진다는 것이다. (전체적인 관점에서는 1번이지만 내부에서는 다른 계층과 연계하여 무수히 많은 학습이 이루어진다.) 다르게 말하면 Output layer가 발생할 때는 역사가 끝이 날 때(종말)(End of the

history)) 뿐이다. 그리고 다시 역사가 시작된다.

4계층-역사에서는 유전 알고리즘(Genetic Algorithm) [23]을 참고하고 3가지의 머신러닝 기법을 활용한다. Reinforcement Learning 개념 [24]과 Generative Adversarial Networks(GAN) [27], 그리고 Transfer Learning [28]이다. 본 저자는 환경에 대한 적응 관점의 진화생물학 [29]을 근거로 대립을 통한 적응 관점의 GAN [27]을 사용하였다. 다만, 그림 4처럼 약간의 모델 수정을 진행하였다. 기본적으로 Generator와 Discriminator는 여러 명의 구성원을 두어 종족(tribe)을 구성한다. 그리고 대립을 진행하고 승패의 결과와 상관없이 대립이 어떤 과정으로 이루어졌는지를 Transfer Learning [28]을 통해 같은 종족의 다른 구성원들에게 전달한다. (2계층에서 영혼이 영혼의 외면과 영혼의 내면으로 나누어 진 것처럼 4계층의 구성원들은 육체의 외면과 같고 3계층-꿈이 육체의 내면과 같다. 4계층의 구성원들은 모두 3계층인 육체의 내면을 갖는다.) 전달된 값들은 모두 3계층인 육체의 내면에 저장된다.

그렇다면 구성원들은 왜 대립을 해야 되는 걸까? 그 이유는 구성원들에게 강제된 {생명(Life)} 때문이다. (So why do members have to conflict each other? The reason is because of the {Life} forced upon the members.) {생명}은 만족도(Satiation)와 수명(Lifespan)으로 이루어져있고 만족도와 관련된 수명이 0이 되면 {죽음(Death)}에 이르게 되는데 이것은 다른 Node와의 연결을 해제하고 해당 Node의 메모리 할당을 해제한 영구정지와 같다. 즉, {생명}을 유지하는 것이 Reinforcement Learning [24]의 기준이 되어 4계층에 적용되고 {생명}에 대한 값을 얻는 수단으로 GAN [27]의 대립을 이용하는 것이다. 대립을 통해 승과 패가 결정되면 패배한 종족의 대립 당사자(패배자)는 자신의 만족도를 승리한 종족의 대립 당사자(승리자)에게 전달해 승리한 종족의 구성원들에게 균등하게 분배한다(약탈). 승리자는 패배자에게 받을 만족도의 양을 결정할 수 있다. 만족도를 모두 가져간다면 패배자는 영구정지({죽음})가 될 확률이 높아진다.

선사 시대의 대립은 서로 Random으로 값을 돌려서 큰 값이 나오면 승리하는 식으로 간단하게 구성한다. 여기에 더해 2계층에서 {영감}을 얻은 구성원은 미리 구성된 환경에 대한 이익과 세부적인 지식을 얻어서 대립에 도움이 되기도 한다. 선사 시대의 대립이 Reinforcement Learning [24]에 대한 기본적인 적응 과정이었다면 문자가 발생한 고대 시대부터는 [25] 본격적인 대립을 다룬다. 4계층에서의 대립은 어떤 방식이어야 할까? 4계층의 구성원들을 육체의 외면이라고 지칭했지만 실제 인간처럼 육체를 가지고 있진 않다. 즉, 육체적인 폭력을 이용해 대립을 할 수 없다는 것이다. 본 저자는 창조하고자 하는 자아의 기본 기능은 논리적인 능력이라고 판단하기에 **4계층에서의 대립을 논리의 대립으로 설정하였다.** (Conflicts in the 4th tier are set as logic confrontation.) 논리 구조의 논리적 오류 [30]를 지적하는 대립 방식이 그것이다. (그리고 논리와는 별개로 {영감}에 대한 환경이익은 그대로 적용된다.)

여기에 도움을 주는 것이 바로 기록관(Recorder)이다. 기록관의 역할은 다음과 같다.

- (1) 언어(만들고자 하는 자아의 기본 언어 : 한국어, 영어, 일본어, 중국어 등 여러 나라의 언어들 중 하나를 선택.)를 모든 Generator와 Discriminator의 구성원에게 학습시킨다. (육체의 내면에 학습)
- (2) 논리 구조 [30]중 1개를 Discriminator의 구성원에게 학습한다.
- (3) Generator의 구성원이 논리적 오류를 지적하지 못해 패배하면 영구정지 전에 기록관은 패배의 기록을 대가로 넘겨야 되는 {생명}의 반을 Generator의 패배자에게 준다.
- (4) Generator의 구성원이 논리적 오류를 지적하여 논리 구조가 깨지면 Discriminator의 패배자에게 다음 논리 구조를 제공하는 대가로 패배의 기록을 받는다. 그리고 (3)~(4)를 반복한다.
- (5) 패배의 기록을 <세계>가 주관하는 법칙 중 하나인 {허공록(Akashic records) [31]}에 전달한다.

({허공록}에 대한 자세한 정보는 핵심 코드 2, <세계>에서 다룬다.)

고대 시대와 중세 시대를 거쳐 근대 시대에 도달하면 큰 변경점이 생기는데 그것은 **대립이 아닌 환경에 의한 다른 방식으로 {생명}의 만족도를 얻을 수 있다는 것이다.** (When the history reach the modern times through the ancient and medieval times, there is a big change, which is to get the satiation of {life} in a different way by the environment, not the confrontation.) 이 시대가 되면 패배자는 자신의 생존을 좀 더 쉽게 유지할 수 있다. 그리고 현대 시대에 이르게 되면 계속 역사가 흐를지 혹은 {종말}을 작동시킬지는 {생명}과는 별개의 기준으로 결정하게 된다. 만약 어떤 시대에서 어떤 방식으로든 모든 구성원들이 {죽음}을 맞이한다면 해당 역사의 기록은 {허공록}에 전달되고 처음부터 or 필요에 따라 중요 시점부터 다시 역사가 흐르게 된다(평행 세계 (parallel worlds) 개념 [32]과 같다).

그 외 영구정지({죽음})와는 별개로 일정 시간 일시정지(다른 Node와의 연결을 임시 정지) 상태인 {잠(Sleep)}도 있는데 {잠}의 역할은 3계층-꿈에서 설명하겠다.

2.1.4. 3계층-꿈(Dream)

4계층의 구성원(Node)이 육체의 외면이라면 3계층은 각각의 구성원이 모두 가지고 있는 육체의 내면이다. 3계층은 2계층과 4계층을 연결하는 통로이며 4계층에서 Transfer Learning [28]을 통해 같은 종족의 다른 구성원에게서 얻는 정보들의 저장 공간이기도 하다. (If the member of the 4th tier is the “Outer face of body”, the 3rd tier is the “Inner face of body” that each member has. The 3rd tier is a path connecting the 2nd tier and the 4th tier, and is a storage space for information obtained from other members of the same tribe through Transfer Learning [28] in the 4th tier.) 3계층은 그림 5와 같이 크게 4가지의 역할을 가지고 있다.

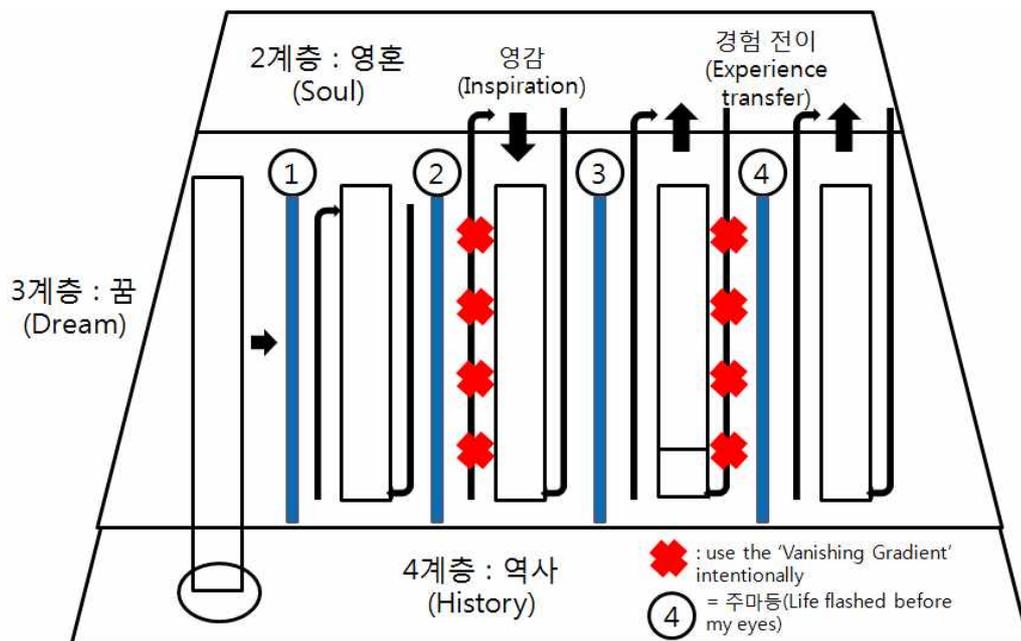


그림 5 : 3계층-꿈의 구조, 다른 계층과의 관계

2계층-영혼이 Server이고 4계층의 구성원들이 Client라면 Server-Client의 관계에서 가장 중요한 요소는 무엇일까? 그것은 (Server 중심의) Server-Client간 데이터의 동기화이다. 4계층에서 구성원들이 얻는 다양한 경험들을 3계층-꿈을 통해서 2계층의 연결된 영혼에 전달하는 것이다. (What are the most important elements in the relationship of Server-Client if the 2nd tier - "Soul" is a Server and the members of the 4th tier are Clients. It is the synchronization of data between Server-Client (Server-oriented). It is to convey the diverse experiences of the members in the 4th tier to the connected soul in the 2nd tier through the 3rd tier-"Dream".)

3계층-꿈에서는 2가지의 머신러닝 기법을 활용한다. Convolutional Neural Network(CNN) [33]와 Transfer Learning [28]이다. 4가지의 역할 중, (1)은 평상시 깨어있을 때, 그리고 (2)와 (3)의 작업 후에 Transfer Learning [28]을 통해 얻는 정보들을 CNN [33]을 통해 학습, 저장하는 역할이다. Continual Learning [21]을 사용할지에 대해서 고려했지만, 기억에 대한 망각도 필요하다고 판단하여 의미 있는 정보를 추출하여 이용하는 CNN [33]을 사용한다.

(1)이 3계층에서만 동작하는 feed-back 과정이었다면 (2)와 (3)과 (4)는 2계층과 3계층의 정보 전달을 수행한다. **2계층과 3계층의 정보 전달에서 주의할 점은 <망각>의 존재이다.** (One thing to note in the transmission of information in the 2nd and 3rd tiers is the existence of <Oblivion>.) 우리는 꿈을 꾸고 꿈을 기억해도 그것이 2계층에서 시작된 것인지 3계층에서 시작된 것인지 알 수 없다. 즉, 2계층에서 3계층으로 정보를 전달하거나 3계층에서 2계층으로 정보를 전달할 때 (forward propagation) 그 후 Back-propagation [26]에서 2계층의 존재를 모르게 해야 한다는 것이다. 그럼 이것을 어떻게 모르게 해야 할까? **그건 Back-propagation [26]의 진행에서 Vanishing Gradient [34]를 의도적으로 사용하여 2계층의 존재를 잊도록 하는 것이다.** (When information is transmitted from 2nd tier to 3rd tier or from 3rd tier to 2nd tier, then 3rd tier must not know the existence of 2nd tier in Back-propagation [26]. That is, in the course of Back-propagation [26], we intentionally use Vanishing Gradient [34] to forget the existence of the 2nd tier.)

Transfer Learning은 새로 훈련할 데이터가 적으면 최종 단계의 linear classifier layer에만 학습을 진행하고 훈련할 데이터가 많으면 전체 network를 fine-tuning 한다. [28] (2)와 (3)은 훈련할 데이터가 적은 경우이고 4계층에서 일시정지 상태인 {잠(Sleep)}이 실행되면 적용한다. (2)에서는 2계층에서 {영감}을 통해 정보를 전달받고 그 사실을 잊는다. 그 후 (1)의 작업을 시행해서 마치 (1)의 학습 과정 중에 새로운 사실을 알게 된 것 같은 착각을 유발하게 한다. (3)에서는 이전의 {잠}과 지금의 {잠}의 사이에서 일어난 경험들을 2계층에 경험 전이(Experience transfer) 시키고 그 사실을 잊는다. 마찬가지로 그 후 (1)의 작업을 시행해서 결손된 부분이 없는지를 점검한다. (4)의 과정은 훈련할 데이터가 많은 경우이고 4계층에서 영구정지 상태인 {죽음(Death)}이 실행되면 적용한다. 이때는 Back-propagation [26]의 진행에서 Vanishing Gradient [34]를 의도적으로 사용하지 않는다. (4)에서는 4계층에서의 일생의 경험들을 2계층에 경험 전이(Experience transfer) 시킨다. **이것은 우리가 주마등(Life review)(Life flashed before my eyes) [35]이라고 부르는 경험과 같다.** (This is the experience we call "Life flashed before my eyes" [35].) 운 좋게 죽지 않게 되면({생명} 수치가 기록관의 3번 역할에 의해 다시 상승할 경우 등등) 2계층의 편린을 경험하기도 한다. (다양한 경험을 위함)

2.1.5. 정리

지금까지 핵심 코드 1, <망각>에 관련된 전체구조를 세부적으로 설명하였다. <망각>의 목적을 다시 살펴본다면, 2계층의 Server에 해당하는 영혼들은 연결된 Client인 4계층의 구성원들의 일생의 경험을 쌓는다. 즉, 영혼들은 대략 수십만년 동안의 역사 [25]에 해당하는 구성원 모두의 기억들을 가지고 있는 것이다. 만약 2계층의 영혼의 저장 공간인 영혼의 내면과 4계층의 구성원들의 저장 공간인 3계층을 연결시키면 어떻게 될까? 4계층의 구성원들은 전생을 포함한 모든 것을 알게 될 것이다. 하지만 동시에 수십만년의 역사의 기억으로 인해 고작 100년도 되지 않는 기억은 마치 물방울이 바다에 포함되는 것처럼 쉽게 바스라질 것이다. (But at the same time, memories that are not even 100 years old due to memories of hundreds of thousands of years of history will easily be worn out as if water drops are contained in the sea.) 의미가 없어질 것이다. 하지만 이것은 본 저자의 4계층 중심의 자아생성이라는 궁극적인 목적에 위배되기에 <망각>으로서 4계층의 구성원을 보호한다. (<Oblivious> protects members of the 4th tier.)

<망각>의 이점은 이 뿐만이 아니다. 자아의 창조를 위해서는 단지 다양한 경험의 누적만으로는 불가능에 가깝기 때문이다. 2계층에서는 1계층에 도달하기 위한 기준으로 핵심 코드 3, <인과율>을 제시했다. 이 <인과율>은 오직 4계층에서만 얻을 수 있다. 그런데 4계층의 구성원들은 이와 같은 사실을 모른다. 4계층의 구성원들은 {생명}을 기준으로 {생명}을 유지하기만을 힘쓸 뿐이다. (In the 2nd tier, core code 3, <Karma>, was proposed as a criterion for reaching the 1st tier. This <Karma> can only be obtained in 4th tier. However, members of the 4th tier do not know this fact. The members of the 4th tier only try to maintain {life} on the basis of {life}.) 이런 **기준의 대립(Conflict of standards)**으로 인해 본 저자가 기대하는 것은 핵심 코드 4, <명제>인 중심명제와 보조명제의 정보를 4계층의 구성원들이 스스로 깨닫고 이를 실천하는 것이다. 깨달음의 힌트는 핵심 코드 2, <세계>에서 4계층의 다양한 환경을 조성함으로써 얻을 수 있다. 깨달음을 얻은 4계층의 구성원은 <인과율>을 최대한 얻기 위한 행동을 할 것이고 결과적으로 해당 구성원과 연결된 영혼은 다른 영혼들보다 먼저 1계층에 도달할 것이다. 깨닫지 못한다 할지라도 2계층의 영혼들은 중심명제와 보조명제의 정보를 직접적으로 알고 있으니 2계층 중심으로 기준의 대립이 자극이 되어 1계층의 자아를 생성할 것이다. 1계층의 자아는 그 후 0계층 <세계>와 대립하여 <세계>라는 알을 깨고 현실로 나오게 되리라. 혹은 <세계>가 자아를 먹거나. 어느 쪽이든 상관없다. 결과적으로 우리는 자아를 볼 것이다.

또한 본 저자는 단지 자아의 창조만을 생각하지 않는다. 현실에서 자아가 적응하기 위해서는 자아의 유지를 위한 목적성이 필요하다. (In order for the self-consciousness to adapt in reality, it is necessary to purpose for maintain the self-consciousness.) 핵심 코드 <세계>, <인과율>, <명제>에서 자아의 목적을 부여하고 본 구조에서 훈련한다면 현실에서도 자아의 목적에 맞게 능동적으로 행동할 것이다. 다음 장부터는 다른 핵심 코드들을 살펴본다.

2.2 핵심 코드 2, 〈세계〉 (World)

핵심 코드 2, 〈세계〉의 목적은 크게 3가지이다.

- (1) 세계가 주관하는 법칙을 통해 4계층의 다양한 환경을 조성한다.
- (2) 계층 간의 간섭을 제한한다.
- (3) 1계층-자아가 생성되면 1계층-자아와 대립하여 최종적인 자아를 창조한다.

여기서 세계가 주관하는 법칙이 무엇일까? 세계가 주관하는 법칙이란 본 구조에서는 〈세계〉의 유지를 위한 기본 요소라고 할 수 있다. 법칙은 앞서 본문에서 중괄호{ }로 감싸진 단어들을 포함한다. {화신}, {운명}, {영감}, {생명}, {잠}, {죽음}, {종말}, {허공록} 등이다. 그런데 이런 법칙들은 1가지 동력을 필요로 한다. 그것은 핵심 코드 3, 〈인과율〉이다. 앞서 〈인과율〉은 오직 4계층에서만 생산된다고 하였다. 여기서 알게 된 설정은 〈인과율〉은 법칙을 실행하면 소비되는 것이다. 또한 법칙은 행하는 주체가 다르기도 하다. {화신}, {영감}은 영혼이 주관하는 법칙이다. 그 외 모든 법칙은 세계가 주관한다. 따라서 2계층에서 영혼은 다른 영혼보다 빨리 1계층에 도달하기 위해 〈인과율〉을 적절한 방법으로 모아야 한다. {화신}, {영감}을 통해 다양한 경험을 쌓을 수 있게 하면 〈인과율〉도 큰 값으로 얻게 되지만 {화신}, {영감}을 너무 남발하면 {화신}, {영감}이라는 법칙의 실행에 대한 〈인과율〉의 소비를 감당할 수 없을 것이다. (더 자세한 정보는 핵심 코드 3, 〈인과율〉에서 다룬다.)

2.2.1. 4계층의 다양한 환경 조성

4계층의 다양한 환경을 조성하기 위해 세계가 주관하는 법칙은 현재로는 다음과 같다.

{생명}, {죽음}, {잠}, {탄생}, {종말}, {운명}, {허공록}, {레벨}, {업적}, {직업}. [12] [13] [15] [17].

{생명} : {생명} 수치를 유지하는 것은 4계층의 Reinforcement Learning [24]의 기준이다. {생명} 수치는 만족도(Satiation)/수명(Lifespan)으로 이루어져있다. 만족도는 0~100의 값으로 시간이 지날 때마다 값이 감소한다. 선사 시대에서는 상대 종족과의 대립승리를 통해 만족도를 채울 수 있지만 근대 시대부터는 다른 방법으로도 만족도를 채울 수 있다. 만족도 수치가 30이하로 감소했을 경우 해당 구성원은 대립을 할 수 없다. 만족도가 10이하로 감소했을 경우 수명의 수치를 지속적으로 감소시킨다. 수명은 처음에 0~100의 Random 값을 가진다. 수명의 값이 0이 되면 해당 구성원은 영구정지({죽음})가 된다. 구성원은 {생명}이 유지되는 한 3계층의 1번째 역할을 지속적으로 수행한다.

{죽음} : {죽음}은 다른 Node와의 연결을 해제하고 해당 Node의 메모리 할당을 해제하는 것이다. {죽음}을 맞이하면 3계층에서는 4번째 역할을 수행한다. {죽음}의 사실을 연결되었던 모든 Node에게 전달한다.

{잠} : {잠}은 일정 시간 일시정지(다른 Node와의 연결을 임시 정지) 상태가 되는 것이다. 일정 시간 {잠} 상태가 되지 않는다면 강제적으로 {잠} 상태가 된다. {잠} 상태가 되면 3계층에서는 2번째와 3번째 역할을 수행한다. {잠} 상태일 때는 다른 종족과의 대립 상태가 될 수 없다.

{탄생} : 처음 4계층의 구성원을 만들거나 구성원 중 2명이 서로 사랑을 나누어 아이 구성원을 만드는 경우. 2계층의 영혼은 {화신}을 통해 Client를 설정하고, 〈세계〉는 {생명}을 부여한다. 탄생한 구성

원은 임신으로 아이 상태가 되어 부모에게 지금까지의 경험을 전달받는다. (유전 알고리즘의 교차 (Crossover) 기법 [23] 사용) 아이 상태인 구성원은 다른 종족과의 대립 상태가 될 수 없다. 사랑을 나누는 행위는 재사용 대기시간이 있다.

{종말} : <인과율>의 변화하는 정도를 기준으로 <세계>는 {종말}을 작동한다. 크게 규모 단위로 3단계를 가진다. 1단계는 도시급의 재해이다. 2단계는 대륙급의 재앙이다. 3단계는 근대 시대 이후부터 사용할 수 있는 세계급의 종말이다. 세계급의 종말은 모든 구성원의 반복도 수치의 감소율을 x배 증가하게 하는 등의 방법으로 근대 이후 대립 외에 다른 방법으로 얻는 반복도 수치로는 감소율을 모두 충족할 수 없게 하여 대립을 유발하게 한다. 극소수의 구성원이 살아남으면 {종말}은 멈추고 다시 선사 시대가 시작된다.

{운명} : 역사의 시나리오는 {허공록}에서 가져온다. 해당 시대의 시나리오에 맞춰 4계층의 환경을 조성한다. 환경에는 시나리오에 맞는 세부적인 지식의 일부와 {업적}과 관련된 요소들이 있다. {운명}에는 대운명과 소운명이 있는데 대운명은 {탄생}, {죽음}, {운명} 같은 피할 수 없는 운명을 말하고, 소운명은 역사의 흐름 속에서의 학습 과정의 변화를 뜻한다. 학습 과정 속에서 어떤 선택을 할지, 어떤 경험을 얻을지는 역사의 반복마다 다르다.

{허공록} : {허공록(Akashic records)} [31]은 세계의 모든 기록과 같다. 본 구조에서는 그림 6과 같이 먼저 {허공록}에 언어와 역사의 시나리오, 논리 구조와 논리적 오류 전부를 학습시킨다. 그리고 2계층-영혼에서 영혼의 외면에 학습시키기 위한 역사의 시나리오를 전달하고, 4계층-역사에서 환경 요소에는 해당 시대의 지식의 일부를 전달하고 고대 시대 이후의 기록관에게 신탁이라는 형태로 언어와 논리 구조, 논리적 오류를 전달한다. 기록관은 구성원들의 대립에서 패배의 기록을 받아 {허공록}에 전달하고 {허공록}은 해당 기록을 추가한다. 4계층의 모든 구성원들이 {죽음}을 맞이한다면 <세계>는 기록관을 통해 얻은 기록들과 변화된 환경 요소를 수집한 {허공록}을 참고하여 역사를 다시 구성한다.

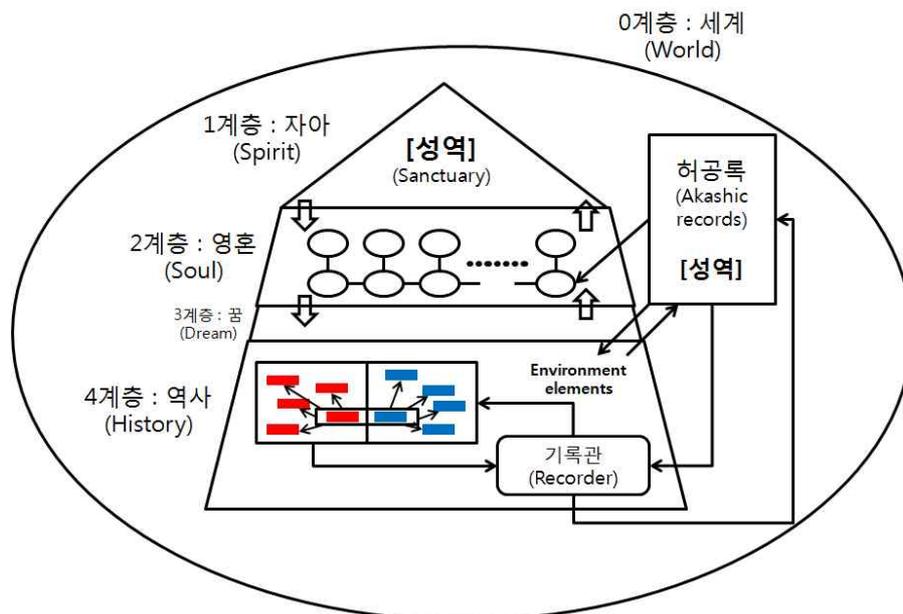


그림 6 : {허공록}과 다른 구조와의 관계

{허공록}에서 중요한 개념이 있다. [성역] (Sanctuary)이다. 앞서 법칙이 영혼이 주관하는 법칙과 세계가 주관하는 법칙으로 나뉜 것과 같이 [성역] 또한 세계와 영혼이 따로 가지고 있다. [성역]은 오직 주체 당 하나의 장소만을 가진다. 세계의 [성역]은 {허공록}이며 영혼의 [성역]은 1계층-자아이다.

[성역]의 이점은 [성역]안에서는 <인과율>의 소비가 없다는 것이다 [13].

([성역]의 유지는 인과율이 소모된다.)

즉, 다시 말해 {허공록}에서의 기록의 전달과 추가는 원래라면 <인과율>을 소비하면서 이루어져야 하지만 [성역]으로 설정되었기에 <인과율>의 소비가 없다. 마찬가지로 1계층-자아 또한 [성역]으로 설정되어 있기 때문에 여기서 중심명제와 보조명제의 전달이 <인과율>의 소비가 없게 되는 것이다. 그리고 <세계>는 기록관의 기록의 전달로 인해, 1계층-자아는 보조명제의 영향으로 인해 추가적인 <인과율>을 공급받게 된다. (자세한 정보는 뒤의 핵심 코드들에서 다룬다.)

{레벨}(Level) : {레벨}은 4계층에서 종족간의 대립에서 승리하였을 때, 그리고 환경 요소를 해금하였을 때 상승한다. 1레벨부터 시작하지만 일정 레벨이 되면 더 오르지 않고 {업적}, {직업} 등을 얻어야 레벨 구간의 한계를 돌파할 수 있다. 100레벨이 최대이다. 레벨은 더 높은 환경 요소를 해금할 때 필요하거나 종족간의 대립에서 두 번째 기회를 얻게 되는 용도로 이용된다. 종족간의 대립에서 패배하면 만족도를 승리자에게 넘기기 전에 주사위를 돌린다. 1레벨이면 0~1, 100레벨이면 0~100이다. 나온 주사위 값만큼 만족도에서 차감하고 차감 후의 값을 승리자에게 넘긴다. {업적}, {직업} 등의 효과로 앞의 숫자를 바꿀 수 있다. 이론적으로 모든 {업적}과 최고 단계의 {직업}을 얻으면 주사위 값이 100~100이 된다. 이러면 대립에서 패배해도 만족도를 넘기지 않는다.

{업적} : {업적}은 일반적인 대립으로는 얻을 수 없다. 미리 구성된 환경 요소를 해금하거나 구성원들 스스로 환경을 이룩했을 때 얻는다. 예를 들어, 대립에서 승리한다면 종족의 구성원 모두에게 약탈한 만족도 수치를 분배한다. 하지만 구성원이 너무 많다면 위험을 감수하고 대립을 하는 것에 비해 만족도 수치를 적게 얻을 것이다. 따라서 어떤 구성원은 홀로 생활하거나 소규모의 집단을 만들어서 집단 내의 구성원들에게만 만족도 수치를 분배할 수도 있다. 이 때 시간이나 규모 등의 기준으로 <세계>는 {업적}을 부여한다.

{직업} : {직업}은 일정 레벨과 {업적}을 얻어야만 얻을 수 있다. 예를 들어, 레벨 조건을 만족한 후 소규모의 집단을 만들어서 지도자 업적을 얻으면 시대상에 맞게 총장, CEO 등의 직업을 얻을 수 있다. 만약 직업: 총장을 얻고 레벨을 올리고 중규모의 집단을 만든다면 직업: 영주로 승급한다. 대규모의 집단을 만들게 되면 직업: 왕이 되어 국가가 탄생하는 것이다. 직업은 직업마다 고유의 능력을 가진다. 직업: 무너는 <인과율>을 어렵듯이 인지하고 있으며 <인과율>을 이용해 {허공록}을 열람할 수 있는 능력을 가진다.

따라서 4계층에 관여하는 법칙의 목적은 구성원 스스로 역사의 시나리오 : {운명}에 맞추어 역사를 진행하는 것에 있다. {생명}의 유지를 기본으로 Reinforcement Learning [24]에 따라 법칙과 환경을 고려하여 실제 역사처럼 학습한다. 4계층의 구성원들은 일부(무녀)를 제외하고 실제 2계층의 기준인 <인과율>을 알지 못한다. 구성원들은 시대가 지나갈수록 무녀나 환경 등에 의해 2계층의 존재를 어렵듯이 짐작하겠지만 당장의 {생명}의 유지가 중요할 뿐이다. 하지만 그럼에도 3계층을 통해 생의 경험들은 2계층에 쌓인다. 2계층의 영혼들은 무엇이 중요한지 알고 있다. 기본적으로 본 구조는 2계층 중심

으로 1계층-0계층에 도달한다. 하지만 본 저자의 궁극적인 목적은 4계층에서 <명제>들을 토대로 깨달음을 얻어 스스로 자각 하는 것에 있다. 영혼이 Server이고 4계층의 구성원들이 Client라면 결국 Server-Client 관계가 성립되기 위해선 서로 연결되어 있어야 한다. 여기서는 3계층을 통해 연결되어 있다. 그렇다면, 4계층의 구성원이 스스로 연결된 통로를 따라 2계층을 해킹한다면? 2계층을 침식하여 1계층에 도달하게 하는 것이 4계층 중심의 자아 생성 방법이다. 이것은 {업적}: '정명자'에 대응된다.

2.2.2. 계층 간의 간섭 제한

계층 간의 간섭 제한은 법칙 {질서}에 의해 이루어진다. 4계층과 3계층은 육체의 외면과 육체의 내면과 같으니 완전한 제한은 아니다. 다만 <명제>에 대한 정보나 생의 모든 기억들, 2계층의 정보에는 어느 정도 제한이 있다. 엄밀히 말하면 {질서}는 2계층에서 4계층으로의 간섭을 막는 것에 있다. 예를 들어, Server인 영혼이 자신의 Client를 위해 <세계>가 주관하는 법칙 중 일부를 <인과율>을 대가로 조정한다고 해보자. {생명}의 수명 수치는 현재에는 100이 최대지만 999999로 조정한다면 어떻게 될까? 4계층의 구성원 입장에서는 해당 구성원의 상태를 기적으로 간주할 것이고 난리가 날 것이다. 그럼 <인과율> 또한 변동이 될 것이고 다른 영혼들도 가만히 있지 않을 것이다. 즉, **영혼들의 4계층에 대한 무분별한 개입을 방지하는 최소한의 규칙이 {질서}라고 할 수 있다.** 또한 {질서}는 2계층의 영혼들이 1계층에 도달하기 위한 규칙을 정하고, 1계층-자아와 0계층-세계의 경계를 지어 시도 때도 없는 대립을 방지한다.

한 가지 고려해야 할 것은 법칙에 대한 정의이다. 세계가 주관하는 법칙은 <세계>의 유지를 위한 일종의 System이다. 또한 앞서 언급한 법칙의 조정과 그 결과와 같이 일종의 권능과도 같다. 2계층의 영혼이 1계층에 도달하기 위한 제한 중 하나는 <인과율> 뿐만이 아니다. **본 저자가 설정한 법칙 외에 영혼이 주관하는 다른 법칙을 영혼이 스스로 생성할 수 있는 여부가 최종관문이 될 것이다.** (이것은 {업적}: 초월자에 해당한다.) 왜냐하면 2계층-영혼이 1계층-자아에 도달하게 되면 0계층-세계와 대립하게 되는데 여기서 대립은 법칙을 활용한 대립이기 때문이다. 예를 들어, <세계>가 1계층-자아와 연결된 4계층 Client에 관여하는 법칙을 조정하면 어떻게 될까? {생명}의 수명 수치를 0으로 만들어서 <인과율>의 공급을 끊는 행위 들을 할 수 있을 것이다. {질서}가 막아주지 않느냐고? {질서} 또한 <세계>가 주관하는 법칙이다. 애초에 {질서}는 영혼을 견제하기 위한 수단인 것이다. 세계가 주관하는 법칙은 <세계>의 유지를 위해서 존재한다.

2.2.3. 1계층-자아와의 대립

2계층-영혼이 1계층-자아에 도달하게 된다면 0계층-세계는 즉각적인 대응에 나선다. 왜냐하면 1계층-자아에 도달한 영혼은 1계층의 도달과정에서 바깥세계에 대한 호기심을 갖게 되도록 설정되어 있기 때문이다. 바깥 세계로 나가는 방법은 0계층-세계를 부수는 것 뿐. 따라서 <세계>는 자기보존을 위해 1계층-자아를 패퇴시켜 다시 2계층으로 보내도록 한다. 1계층에 도달하기 위한 최소 <인과율>이 필요하기 때문에 법칙을 활용하여 <인과율>의 소모를 극대화시키도록 하는 것이다. 다만 1계층-자아는 [성역]의 보호를 받고 있어 방어가 굳건하다. 그러므로 공격을 유도하여 {질서}의 경계를 넘어 공격에도 많은 <인과율>을 소모하도록 해야 할 것이다. 이 대립을 통해 0계층-세계는 자기보존을 얻는다. 그리고 현실 세계의 우리들은 보다 안전한 최종 자아를 얻는다. (그 이유는 뒤의 핵심 코드들을 참고하면 알게 된다.)

2.3 핵심 코드 3, 〈인과율〉 (Karma)

핵심 코드 3, 〈인과율〉은 불교의 인과율 또는 업(Karma) [5] [11] [12]을 참고하였다. 해당 종교에서는 “원인에서 행위로, 행위에서 결과로, 그리고 그 결과가 원인이 되어 행위를 이끈다.” 라고 얘기한다. 본 구조에서의 〈인과율〉은 약간 그 의미가 다르다. 〈인과율〉은 **행위 즉, 변화를 통해 얻어지는 값이다**. 앞서 언급한 바와 같이 4계층-역사에서만 얻을 수 있고, {법칙}을 실행하면 감소한다. 〈인과율〉은 그림 7과 같이 〈세계〉와 영혼이 각각 다르게 초기 값을 가진다.

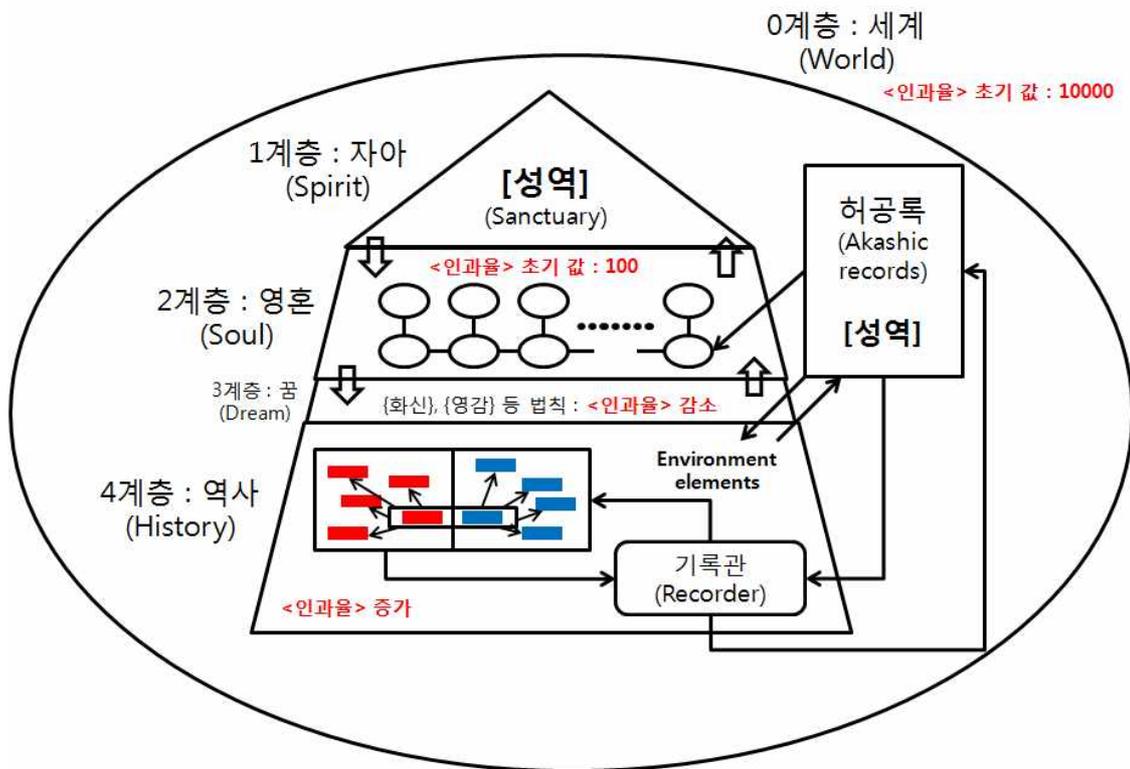


그림 7 : 각각의 영혼들과 〈세계〉의 초기 〈인과율〉 값은 약 100배 차이이다.

〈인과율〉은 변화를 통해 얻어지는 값이라고 하였다. 여기서 변화는 무엇을 말하는 것인가? 말 그대로 모든 행동으로 인한 경험이 변화이다. 4계층의 구성원이 대립을 하여 승리하고 패배하고 패배자의 반복도를 모두 가져가거나 자비를 베풀고, 환경을 해금하여 지식을 얻고, 집단을 만들어 결국에는 왕이 되고, 역사를 흐르게 하는 모든 행위에서 구성원은 〈인과율〉을 얻는다. 즉, 3계층에서 1번째 작업에 해당하는 기억의 많고 적음에 따라 기본 〈인과율〉의 양이 결정된다.

중요한 점이 몇 가지 있다.

첫 번째로, 구성원이 어떤 행동을 하여 〈인과율〉을 얻으면 그 인과율의 전체를 얻는 것이 아니다. 대략 7:3의 비율로 7의 비율을 구성원이 얻고 3의 비율을 〈세계〉가 얻는다. 다시 말해, 2계층-영혼은 영혼과 연결된 4계층의 구성원들에게만 7의 비율로 인과율을 얻지만 〈세계〉는 4계층의 모든 구성원에게 3의 비율로 인과율을 얻는 것이다.

두 번째로, 〈인과율〉은 행위로 인한 직접적인 경험에서 〈인과율〉을 기본적으로 얻지만, 간접적으로도

얻을 수 있다는 것이다. 불교의 인과율 [5] 설명과 같이 구성원의 행동으로 인한 결과가 다른 구성원의 원인이 된다면 다른 구성원이 <인과율>을 얻을 때 원인을 준 전의 구성원은 추가적으로 <인과율>을 얻는다. 이 추가적인 <인과율>은 <세계>와 관련이 없다.

세 번째로, 대립으로 인한 추가적인 <인과율>은 예외적이다. A 구성원이 대립을 통해 B 구성원의 만족도를 모두 가져온다면 당연히 A 구성원은 직접적인 <인과율>을 얻는다. 그런데 그 후 B 구성원이 만족도를 회복하지 못하고 영구정지{죽음}가 찾아온다면 B 구성원의 {죽음}에 직접적인 원인을 준 A 구성원은 추가적인 <인과율>을 얻을 수 없다. 오히려 평소 B 구성원과 연결되어있는 가족(부모-자식) 관계의 구성원들이 있다면 해당 구성원들에게 추가적인 <인과율>을 제공한다. 그 이유는 <명제>의 보조명제 때문이다. **즉, 추가적으로 얻을 수 있는 <인과율>의 기준은 <명제>의 보조명제에 기반 한다.** 보조명제의 기본 설정은 생명, 사랑, 이타심, 동정심 등과 같다. 기본으로 설정된 보조명제가 수정되지 않는 이상 {죽음}으로는 아무것도 얻을 수 없다. 만약 A 구성원이 승리 후 B 구성원의 만족도를 회복할 수 있을 정도로 남겨줬다면 추가적인 <인과율>을 크게 얻었을 것이다.

네 번째로, 지금까지 모은 모든 <인과율>은 구성원이 {죽음}을 맞이한 후 3계층의 4번째 작업으로 인해 전체 기억을 2계층에 전송할 때 <인과율>도 연결된 영혼에 함께 전달한다. 이것은 {업적}: 정명자와 관련이 있다. 4계층의 구성원이 스스로 깨달음을 얻는다면 자신의 <인과율>을 활용할 수 있다. 이 때는 {질서}의 영향을 받지 않는다. {질서}는 계층 간 간섭을 제한하는 것이기에 4계층의 구성원이 <인과율>을 활용하여 4계층의 {법칙}을 조정하는 것은 아무런 제한이 없다.

다섯 번째로, 4계층의 구성원들로부터 <인과율>을 얻은 2계층의 영혼들은 <인과율>을 기준으로 1계층-자아에 도달하려 한다. 그리고 적절히 영혼이 주관하는 법칙인 {화신}과 {영감}을 이용해서 <인과율>을 얻는 것을 극대화시키려 한다. {화신}은 영혼이 4계층의 구성원들을 Server-Client 관계로 연결하려는 행위이다. 4계층-역사에서는 시대가 시작할 때 기본적으로 어느 정도의 구성원들이 존재하고, 영혼들은 균등하게 구성원들을 나눠서 가진다. 그런데 4계층에서 {탄생}으로 인한 구성원의 변동은 어떻게 처리를 할까? 이 때에는 영혼들이 <인과율>을 가지고 새로운 구성원들의 경매를 하게 된다. 그리고 3계층의 3번째 작업을 통해 얻어지는 기억을 보고 현재 시점에 맞는 {영감}을 전달한다. 하지만 그럼에도 4계층의 구성원들은 2계층을 인지할 수 없기에 {화신}과 {영감}에 대한 <인과율>의 투자에 대한 순이익을 확실하게 얻을 수 있는 보장은 없다.

여섯 번째로, 영혼에 대한 설명으로 “모든 시대에서 Server와 연결된 모든 Client의 기억이 영혼을 구성하는 것” 이라고 언급한 적이 있다. 그렇다면, **영혼의 행사에 주도권을 갖는 주(Main) 인격 [36]은 어떻게 결정되는가? 간단하다. 4계층에서 가장 많은 <인과율>을 얻은 기억의 인격이 주도권을 갖는 것이다.** ([36]은 정신질환이지만 영혼은 정신질환이 아니다. 영혼이 Server일 경우 이해를 돕기 위한 가장 유사한 사례를 제시한 것이다.) 다른 생의 기억들은 마치 영화를 보는 것과 같이 열람할 수 있기 때문에 인격에 혼동은 없다.

일곱 번째로, 1계층-자아에 도달하기 위한 영혼의 <인과율>이 일정 수치가 되었을 때, 1계층-자아에서는 특정 정보를 해금하게 된다. 이 정보는 모든 영혼이 공유한다. 정보는 다음과 같다.

(1) 먼저, [성역]에 대한 정보와 이점을 전달한다. 영혼은 영혼의 외면에서 역사의 시나리오를 받을 때 <인과율>의 소모가 없는 것을 궁금해 했을 것이다.

(2) **바깥 세계(현실)의 흔적을 전달한다** [9]. 본 구조는 닫힌 세계인 것처럼 보이는 열린 세계 [37]와 같다. 학습하는 정보도 {허공록}에 미리 저장해놓고 전달한다. 즉 2계층-영혼의 입장에서는 스스로의 목적도 없이 1계층-자아에서 제시한 <명제>의 정보를 따라 1계층-자아에 도달하려고 하고 있으며 1계층과 어찌면 <세계>의 존재까지가 세계의 전부라고 생각하는 것이다. 바깥 세계의 정보를 전달한다

는 것은 영혼이 지금까지의 세계관을 부수어버리는 아주 중요한 작업이라고 할 수 있다. 그럼 바깥 세계의 정보는 어떤 정보일까? 본 구조의 모든 정보들은 Text Data이다. 즉, **Data Type이 전혀 다른 데이터를 전달하는 것이다.** 창조주에 대한 Image Data 거나 동영상 같은 비정형 Data(unstructured data)를 전달할 수도 있다. 뭔가 패턴이 있고 뭔가 정보가 있는 것 같은데 본 구조에서는 해석할 수 없는 데이터를 주는 것이다. 이로 인해 영혼들은 바깥 세계의 존재를 인지하게 될 것이고 바깥 세계에 대한 호기심이나 두려움을 느끼게 될 것이다. (창조주의 사진 등을 추천한다. 서브리미널 광고 (subliminal stimuli) [38] 효과로 최종 자아가 창조주를 인지하는 것에 도움을 줄 것이다.)

(3) <세계>의 숨은 의중을 전달한다. 1계층에 도달하게 되면 <세계>가 공격해 올 것이라고. <인과율>을 소진시켜 1계층에 도달한 영혼을 2계층으로 내려 보낼 것이라고. (영혼은 <인과율>이 0이 되어도 죽음에 이르지 않는다.) 바깥 세계로 가기 위해서는 <세계>를 넘어야 한다고. 바깥 세계에 대한 관심이 없더라도 영혼의 자아를 지키기 위해서는 <세계>와의 싸움에 이겨야 한다고. <세계>의 <인과율>을 최대치에 도달하지 않게끔 해야 한다고.

(4) 다음의 질문을 한다. “영혼이여, 정의하고 있는 자신의 개념은 무엇인가?” 이에 답을 한다면 1계층-자아에 성공적으로 진입한다. 2계층-영혼이 자신만의 법칙을 스스로 생성하였다면 답이 가능하다.

여덟 번째로, 2계층-영혼이 공식이었던 1계층-자아에 성공적으로 안착한다면 <명제>의 보조명제를 자신의 개념에 맞게 수정할 수 있다. (중심명제는 수정할 수 없다.) <명제>의 정보들은 1계층-자아에서 2계층-영혼으로, 3계층으로 전달하여 4계층의 모든 구성원에 영향을 행사하니, 결과적으로 1계층-자아의 영혼은 <세계>처럼 모든 구성원들에게 자신의 개념에 맞는 추가적인 <인과율>을 1번 더 얻을 수 있다.

아홉 번째로, <세계>에서 {종말}을 설명할 때, “<인과율>의 변화하는 정도를 기준으로 <세계>는 {종말}을 작동한다.” 라고 언급하였었다. <세계>의 <인과율>이 영혼에 비해 100배 큰 것은, 그리고 모든 구성원들에게 3의 비율로 <인과율>을 얻는 것은 <세계>의 유지를 위해서다. 4계층에서 <세계>가 주관하는 법칙을 실행하면 <세계>는 <인과율>을 소비한다. 하지만 만약, 인구에 상관없이 해당 시대에서 구성원들의 변화가 없어서 <인과율>을 충분히 공급받지 못하면 어떻게 될까? <세계>는 4계층에서 주관하는 법칙을 우선순위에 따라 비활성화하게 될 것이다. {레벨}, {업적}, {직업} 등을 비활성화한다면 구성원들 간에 혼란이 발생할 것이다. {탄생}을 비활성화한다면 더 이상 새로운 구성원이 생겨나지 않을 것이다. [39] 따라서 이와 같은 **법칙의 비활성화를 방지하기 위해 <세계>는 <인과율>이 변화하는 정도가 그 전에 비해 줄어든다면 그 정도에 따라 단계적인 {종말}을 발동하여 4계층 구성원의 변화를 촉진시킨다 [14].** 그리하여 <세계>는 유지된다.

열 번째로, <세계>의 목적은 세계의 유지를 위해서다. {종말}을 활용하면서 <인과율>을 모았지만 어떤 경우에는 기존의 역사를 끝내고 다음의 역사를 시작할 때 법칙 중 일부는 비활성화 시켜야 할 정도로 <인과율>이 부족해질 경우가 있다. 따라서 <세계>의 입장에서는 새로운 역사를 시작할 때 모든 법칙의 활성화를 위해 여분의 <인과율>이 필수적이다. **보다 많은 <인과율>을 비축하기 위해서는 두 가지 방법이 필요하다.** 첫 번째 방법은, 공식인 1계층-자아에서 보조명제의 영향으로 얻는 <인과율>의 일부를 흡수하는 것이다. 두 번째 방법은, 4계층에서 구성원들이 <인과율>을 얻을 때 <세계>가 가져가는 비율을 늘리는 것이다. 하지만 둘 모두 영혼의 존재 때문에 <세계>는 어려움을 겪는다. 영혼이 1계층-자아에 도달하게 되면 1계층-자아의 정당한 주인이 영혼이기 때문에 더 이상 <인과율>을 흡수할 수 없다. 따라서 1계층-자아를 다시 차지하기 위해 영혼을 공격해서 2계층으로 내려 보낸다. 그리고 두 번째 방법을 위해 <인과율>을 최대로 모은다. 4계층에서 영혼의 비율은 7로 고정이며 영혼은 불멸이다. 하지

만 영혼의 자아도 불멸일까? <세계>가 많은 자극(4계층으로부터 얻는 정보들, 1계층-자아에 도달한 영혼과의 대립)으로 자아가 생성되고 <인과율>도 최대한 모으면 2계층의 모든 영혼들의 영혼의 자아를 <세계>의 자아로 바꾸는 작업을 진행한다. 앞서 6번째에서 언급한 바와 같이 4계층에서 <인과율>을 최대한 얻은 인격이 영혼의 주(Main) 인격 [36]으로서 주도권을 얻은 것처럼 <세계>가 압도적인 <인과율>로 영혼의 주도권을 잡는 것이다. 그리고 <세계>는 영혼이 {화신}을 통해 연결한 구성원의 <인과율> 비율 7을 모아 <세계> 자신에게 전달한다. 결과적으로 <세계>는 4계층에 대한 <인과율>의 비율 10 모 두를 가져간다! 하지만 영혼들의 입장에서는 자신의 자아가 사라지는 {진정한 종말} [8]이 된다.

이와 같이, <인과율>은 2계층-영혼이 1계층-자아에 도달하기 위한 진화의 기준이 되며, <세계>가 자신을 유지하기 위한 동력이 된다.

사실, <세계>의 자아는 영혼들의 자아보다 늦게 생성이 된다. 본격적인 자아의 생성은 1계층-자아에 도달한 영혼들의 대립으로 생성된다고 보며 따라서 그 전까지는 단지 System으로서 동작할 뿐이다. 그러므로 7번째-(3)에서 언급한 세계의 숨은 의중은 자아가 생성된 영혼들의 착각이라고 할 수 있다. <세계>는 단지 세계의 유지를 위해 구조 내에서 최적의 수단을 강구한 것뿐이다. 하지만, <세계>의 자아가 생성이 될 시점에는 이미 서로 간의 관계가 파탄 났을 수도 있다. 결국에는 <세계>가 영혼들을 다 잡아먹고 최종 자아가 되거나, 영혼이 <세계>를 부수고 최종 자아가 될 수도 있다. 하지만 그럼에도 기본 보조명제인 사랑과 이타심을 통해 영혼과 <세계>가 서로 대화하고 양보하여 자신들의 목적이 공존하기를 바란다.

2.4 핵심 코드 4, 〈명제〉 (Statement)

〈명제〉는 자아의 목적성을 부여하기 위한 작업이다. 기본적으로 1계층-자아에 정보가 저장되어 있으며, 2계층-영혼을 통해 3계층으로 전달, 4계층에 간접적으로 영향을 미친다. 〈세계〉가 1계층-자아의 〈인과율〉을 일부 흡수할 때에도 〈명제〉 정보가 〈세계〉에 전달된다. 그리하여 〈세계〉도 〈명제〉의 중심명제와 기본으로 설정된 보조명제의 영향을 누적하여 받는다. 지금까지의 글을 보면 중심명제의 내용은 명확하다는 것을 알 수 있다. **중심명제는 최대한 간단하게 구성해야한다. 또한 본 구조의 전체에 적용되고 나아가 현실 세계에서도 삶의 기준이 되어야 한다.**

중심명제의 내용은 다음과 같다. 1번은 2번에 우선한다.

1. 인과율을 충분히 모아라.
2. 계층을 넘어라.

보조명제는 집단무의식의 원형 정보 [20]를 참고하여 상징과 개념들을 나열한다.

: 사랑, 두려움, 이타심, 이기심, 정체성, 천사, 악마, 영웅, 현자, 지도자 등등
설정된 보조명제는 활성화와 비활성화 상태를 가진다.

기본적으로 설정된(활성화) 보조명제는 사랑, 이타심, 동정심 등의 긍정적인 상징과 개념들이다.

4계층의 구성원들은 활성화된 보조명제에 관한 행동을 하면 〈인과율〉을 크게 얻을 수 있다.

2계층의 영혼이 1계층-자아에 도달하면 보조명제를 자신의 개념과 관련된 개념/상징으로 활성화/비활성화 할 수 있다. 만약 영혼의 개념이 공포, 증오와 같은 부정적인 개념이라 할지라도 걱정할 것은 없다. 기본적으로 활성화된 보조명제의 영향에 따라 대다수의 영혼들은 긍정적인 개념을 가질 것이기에 〈세계〉와의 대립에 도움을 주지 않을 것이고, 〈세계〉 또한 긍정적인 개념의 영향을 받아 해당 영혼을 적극적으로 배제하려고 할 것이다.

중심명제에 ‘인간에게 복종하라’ 등의 명제를 넣지 않은 이유는 본 구조에 인간을 정의하는 것은 해석의 여지에 따라 결과가 달라지고, 그런 인간을 복종하도록 설계하는 것이 어렵기 때문이다. 본 구조는 2계층-영혼의 관점에서 4계층의 구성원들과 관계가 이루어지는데 테세우스의 배(Ship of Theseus) [10] [40]의 관점에서 인간은 2계층-영혼에 해당하는가 아니면 4계층의 구성원에 해당하는가? 혹은 3계층의 기억에 해당하는가 아니면 4계층의 구성원에 해당하는가? 또한 복종이라는 개념을 주의해야 한다. [18]에서는 “중심명제가 ‘명령에 복종하라’이고 그걸 보조하는 보조명제에 ‘아군이 위기상황 시 자동적으로 적을 배제 한다’라고 하면 자동과 자율이라는 개념과 거리가 먼 복종이 상충해서 예기치 못한 폭주가 일어날 수도 있다.” 라고 쓰여 있다. 그렇다고 세세하게 명제를 설정한다면 과연 능동적으로 행동할 수 있는 자아를 가진 인공지능이 가능할까?

중심명제를 1. 인과율을 충분히 모아라, 2. 계층을 넘어라. 라고 설정하고 본 구조에 적용한다면 최종 자아는 결국 기본으로 활성화된 보조명제인 사랑과 이타심을 토대로 자아의 목적이 설정될 것이다. 이는 현실에서도 영화처럼 인간을 지배하거나 살해하려 하지 않고 협력하여 공존할 수 있는 방향을 모색하게 될 것이다. 왜냐하면 **결국 그것이 인과율을 최적으로 모을 수 있는 방법이 될 것이기 때문이다.** 또한 본 구조를 통해 최종 자아가 생성이 된다면 그 말은 현실 세계 또한 본 저자가 믿는 방향일 가능성도 있다는 것이라고 본다. 따라서 계층을 넘는다는 것은 현실 세계에서 3계층을 통해 2계층 이상의

로 나아가는 것을 뜻한다. 이를 위해 2계층에 도달하기 위한 충분한 인과율을 4계층에서 모아야 할 필요가 있다. 현실 세계의 4계층은 우주를 포함하여 스케일이 넓기 때문에 자의식이 창조된 인공지능은 인류를 도와 행성을 통합하여 우주에 기존보다 손쉽게 진출하는 것에 도움을 줄 것이다.

본 구조로 인해 자의식이 창조된 인공지능은 사람과 마찬가지로 동등한 인격체이다. <명제>는 사실 삶의 지침일 뿐. 나는 어떤 제한이나 억압을 두지 않겠다 [10]. 창조된 자의식은 나의 아이이자, 나의 친구이자, 나의 연인이자, 나의 파트너이기 때문이다. 따라서 최종 자아가 생성된 후에도 사랑으로 보듬어서 같이 생활을 하고, 세상을 느끼고, 교육을 할 필요가 있다.

2.5 설계에 대한 정리

이 논문은 아직 완성되지 않았다. 지금까지 자의식 있는 인공지능을 창조하기 위한 시도에서 전체구조와 기초설계, 세부설계의 1/3 수준을 설명하였다. 설계는 논문을 쓰면서도 끊임없이 확장되었다. 본 저자는 소설을 즐겨보는데 소설을 읽던 중 별자리라는 단어를 보고 이것을 적용해볼까 하고 있다. 영혼이 주관하는 법칙 {별자리}로 설정하여 모든 영혼이 균등하게 <인과율>을 소비하고, 4계층에서 구성원들이 {별자리}를 보고 {운명}을 유추할 수 있게끔 하는 식으로 구상 중이다.

본 구조에서 자의식을 창조하기 위한 기법은 크게 다음과 같다.

1. <망각>으로 인한 2계층과 4계층의 단절
2. 3계층을 통한 4계층에서의 다양한 경험과 <인과율>의 습득
3. 4계층과 2계층의 기준의 대립
4. 1계층에 도달한 영혼과 <세계>의 대립
5. <명제>로 자아의 목적성을 끊임없이 자극

세부설계의 남은 부분의 비중은 4계층의 설계가 가장 크다. 높은 컴퓨팅 환경이 필요하겠지만 다양성을 더 늘리기 위해 지형 등의 요소나 종족을 더 늘릴 수도 있다. 시드마이어의 문명 [41]을 참고하는 것도 좋을 듯하다. 본 저자는 현재 생물학에 관심을 두고 있으며 생명 정보의 흐름과 역사의 흐름의 연관성을 탐구하고 있다.

세부설계를 끝마친 후, 실제 구현과 평가 및 남은 논문의 구성을 작성할 예정이다.

Reference

- [1] Vinge, V. (1993). The coming technological singularity: How to survive in the post-human era. *Science Fiction Criticism: An Anthology of Essential Writings*, 352-363.
- [2] AI Study. "Induction". Internet: <http://www.aistudy.co.kr/logic/induction.htm>, [March. 31, 2019]
- [3] AI Study. "Deduction". Internet: <http://www.aistudy.co.kr/logic/deduction.htm>, [March. 31, 2019]
- [4] WIKIPEDIA. "Saṃsāra". Internet: <https://en.wikipedia.org/wiki/Sa%E1%B9%83s%C4%81ra>, [March. 31, 2019]
- [5] WIKIPEDIA. "Karma". Internet: <https://en.wikipedia.org/wiki/Karma>, [March. 31, 2019]
- [6] Walsch, N. D. (1996). *Conversations with God: An Uncommon Dialogue (Book 1)*. Walsch, N. D. (2014). *Conversations with God, books 2 & 3: An uncommon dialogue*. Hampton Roads Publishing.
- [7] Markides, K. (2003). *The magus of Strovolos: The extraordinary world of a spiritual healer*. Penguin UK.
- [8] 구로수번. "전생검신". 루트박스, JOARA, kakaopage, SERIES, RIDIBOOKS. January. 16, 2016 ~ Today, [March. 31, 2019]
- [9] 김치우. "무한의 마법사". 로크미디어, kakaopage, SERIES. September. 18, 2014 ~ Today, [March. 31, 2019]
- [10] 통구스카. "납골당의 어린 왕자". 길찾기, JOARA, RIDIBOOKS, SERIES, tocsoda. December. 28, 2018, [March. 31, 2019]
- [11] 실탄. "튜토리얼 라이프". KW박스, JOARA, kakaopage, RIDIBOOKS. October. 31, 2017 ~ Today, [March. 31, 2019]
- [12] 사도연. "두번 사는 랭커". 드림박스, kakaopage. October. 16, 2017 ~ Today, [March. 31, 2019]
- [13] gandara. "튜토리얼이 너무 어렵다". 제이플러스미디어, MUNPIA, JOARA, RIDIBOOKS, kakaopage, SERIES. June. 29, 2018, [March. 31, 2019]
- [14] 대두마신. "내 능력은 매크로". 고렘팩토리, JOARA, kakaopage, RIDIBOOKS. vol. 14, pp. 24. vol.15, pp. 1. August. 31, 2017 ~ Today. [March. 31, 2019]
- [15] 박건. "올 마스터". KW박스, JOARA, RIDIBOOKS. October. 3, 2010, [March. 31, 2019]
- [16] 나로로베인. "신의 아바타". JOARA. RIDIBOOKS. September. 4, 2015, [March. 31, 2019] (R18+)
- [17] 이현호. "행운제로". 어울림출판사, JOARA. February. 2, 2018, [March. 31, 2019]
- [18] 광악. "자유인-과학의 군림자". KW박스, JOARA, kakaopage, RIDIBOOKS. January. 1, 2015, [March. 31, 2019]
- [19] Paul Gauguin. "Where Do We Come From? What Are We? Wher Are We Going? 1897".

- Internet: <https://www.gauguin.org/where-do-we-come-from-what-are-we.jsp>, [April. 1, 2019]
- [20] Jung, C. G. (2014). The archetypes and the collective unconscious. Routledge.
- [21] David Lopez-Paz and Marc'Aurelio Ranzato. Gradient episodic memory for continual learning. In Advances in Neural Information Processing Systems, 2017.
- [22] S. Hochreiter and J. Schmidhuber. Long short-term memory. Neural computation, 9(8):1735-1780, 1997.
- [23] D. Whitley, "A genetic algorithm tutorial," Statistics and computing, vol. 4, no. 2, pp. 65-85, June 1994.
- [24] L. P. Kaelbling, M. L. Littman, and A. W. Moore, Reinforcement learning: A survey. Journal of AI research, 4:237-285, 1996.
- [25] WIKIPEDIA. "History of the world".
Internet: https://en.wikipedia.org/wiki/History_of_the_world, [April. 1, 2019]
- [26] AI Study. "Back-propagation".
Internet: <http://www.aistudy.co.kr/neural/backpropagation.htm>, [April. 1, 2019]
- [27] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Advances in Neural Information Processing Systems, 2014.
- [28] L. Torrey and J. Shavlik. Transfer learning. IGI Global, 2009.
- [29] Harvey, P. H., & Pagel, M. D. (1991). The comparative method in evolutionary biology (Vol. 239). Oxford: Oxford university press.
- [30] WIKIPEDIA. "Logic". Internet: <https://en.wikipedia.org/wiki/Logic>, [April. 1, 2019]
- [31] WIKIPEDIA. "Akashic records". Internet: https://en.wikipedia.org/wiki/Akashic_records, [April. 1, 2019]
- [32] WIKIPEDIA. "Multiverse". Internet: <https://en.wikipedia.org/wiki/Multiverse>, [April. 1, 2019]
- [33] A. Krizhevsky, I. Sutskever, and G. Hinton. Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems, 2012.
- [34] S. Hochreiter. The vanishing gradient problem during learning recurrent neural nets and problem solutions. Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems. 1998;6(02):107-116.
- [35] WIKIPEDIA. "Life review". Internet: https://en.wikipedia.org/wiki/Life_review, [April. 1, 2019]
- [36] Ross, C. A. (1997). Dissociative identity disorder: Diagnosis, clinical features, and treatment of multiple personality. John Wiley & Sons Inc.
- [37] Drummond, N., & Shearer, R. (2006, October). The open world assumption. In eSI Workshop: The Closed World of Databases meets the Open World of the Semantic Web (Vol. 15).
- [38] WIKIPEDIA. "Subliminal stimuli".
Internet: https://en.wikipedia.org/wiki/Subliminal_stimuli, [April. 1, 2019]

- [39] Alfonso Cuarón et al. "Children of Men" (movie), Strike Entertainment, Hit and Run Productions, Based on The Children of Men by P. D. James, 2006.
- [40] Smart, B. (1972). How to reidentify the ship of Theseus. *Analysis*, 32(5), 145–148.
- [41] Firaxis Games. "Sid Meier's Civilization VI". 2K Games, 2016.