JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 25, No. 2, pp. 431-439, Feb. 2024

초등학생 도덕성 발달에 영향을 미치는 모럴 머신 프로그램 개발

문 수 지[1] · 김 성 백[2*]
[1]제주대학교 교육대학원 인공지능융합교육전공 석사과정
[2*]제주대학교 교육대학원 인공지능융합교육전공 교수

Development of a Moral Machine Program Influencing the Moral Development of Elementary School Students

Su Ji Moon[1] · Seong Baeg Kim[2*]

[1]Master's Course, AI Convergence Education Major, Jeju National University, Jeju 63243, Korea
[2*]Professor, AI Convergence Education Major, Jeju National University, Jeju 63243, Korea

[요 약]

최근 사회 전반에 걸쳐 인공지능이 활용됨에 따라 교육에서도 인공지능을 활용한 교육이 관심을 끌고 있다. 비윤리적인 출력을 내는 인공지능에 대한 경계가 높아지고 인공지능 윤리 교육을 제안하는 연구 또한 활발해지고 있다. 본 연구에서는 인공지능 윤리 교육 중 '윤리적인 AI'에 대한 교육에 초점을 맞춰 14차시의 프로그램을 개발하였다. AI에 관심이 있는 A지역 초등학교 5, 6학년 학생들을 대상으로 한 본 연구는 학습자 실태 분석을 위한 사전 설문지와 도덕성 함양을 확인하기 위한 도덕성(공정) 사전 사후 검사, 산출물 수행평가 등을 실시하였다. 검사 분석 결과 학생들은 '공정'이라는 가치를 더욱 깊게 이해하고 내적으로 함양하고 있었고 엔트리를 활용한 프로그래밍 능력도 전보다 신장되었음이 확인되었다. 본 연구에서 개발한 프로그램은 학생들의 도덕성 함양과 인공지능에 대한 이해 및 기초적인 프로그래밍 능력 향상을 위한 자료로 활용될 수 있을 것이다.

[Abstract]

As artificial intelligence (AI) has been utilized across the society in recent years, it has also attracted attention in the field of education. There is growing concern about AI producing unethical outputs, and researchers are proposing AI ethics education. Vigilance against AI that produces unethical output is increasing, and an increasing amount of research is recommending AI ethics education. In this study, we developed a program for the 14th session by focusing on "ethical AI" in AI ethics education. This study targeted fifth and sixth grade elementary school students in 00 city who were interested in AI. Moreover, it conducted a preliminary questionnaire survey to analyze learner status, a program satisfaction survey, and a product performance evaluation. The results confirmed that students had a deeper understanding of and internally cultivated the value of "fairness," and their coding ability using entries had also improved compared with the previous level.

색인어 : 공정성, 모럴 머신, 도덕적 딜레마, 윤리적인 AI, 회귀분석
Keyword : Fairness, Moral Machine, Moral Dilemma, Ethical AI, Regression Analysis

http://dx.doi.org/10.9728/dcs.2024.25.2.431

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-CommercialLicense(http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 21 December 2023; **Revised** 01 February 2024
Accepted 19 February 2024

***Corresponding Author; Seong Baeg Kim**
Tel: +※ 개인정보 표시제한
E-mail: sbkim@jejunu.ac.kr

Ⅰ. 서 론

1-1 인공지능 윤리 - 윤리적인 AI(Artificial Intelligence)

인공지능의 활용은 날이 갈수록 사회 각 분야에서 활발하게 이루어지고 있다. 이에 정부 및 교육 현장에서 인공지능 교육에 대한 관심과 수요가 끊임없이 증가하고 있다. 교육부는 이러한 변화에 발맞춰 교육과정을 개정하고 2022 개정교육과정 총론에서 개정의 이유를 “인공지능 기술 발전에 따른 디지털 전환 (중략)에 의해 사회의 불확실성이 증가하고 있다.”[1]라고 발표하였다. 2022 개정교육과정에서는 기존 초등학교 실과 교과 내용 체계에 ‘디지털 사회와 인공지능’이라는 영역을 신설하여 중학교 1~3학년 정보 교과와 연계되어 변화하는 세상을 인식하고 컴퓨팅 사고를 바탕으로 인공지능을 활용한 실생활 문제해결력 함양을 지향한다[2].

인공지능 활용이 증가함에 따라 최근 비윤리적인 출력을 내는 인공지능의 위험성에 대한 인식이 높아지고 있다. 특히 2015년 발생한 구글포토의 흑인 비하 ‘고릴라 사건’, 2016년 마이크로소프트(MS)사의 챗봇 ‘TAY’의 유대인 차별 발언 사건뿐만 아니라 국내 스캐터랩에서 개발한 챗봇 ‘이루다 사건’ 등은 우리에게 인공지능 윤리 교육이 얼마나 중요한지 보여주고 있는 대표적인 사례이다. 초등학교 도덕 교과에서는 인공지능과 관련된 윤리가 따로 다루어지지 않지만, 인공지능 윤리 교육은 우리에게 당면한 과제이며 타 교과와 융합하여 인공지능 윤리 교육을 제안하는 연구들이 많이 이루어지고 있다.

“인공지능 윤리교육 안에 큰 두 항목인 AI의 윤리와 윤리적인 AI가 있다.”[3] 그러나 현재 ‘AI의 윤리’ 교육에 대한 제안 연구는 활발하지만, ‘윤리적인 AI’에 대한 교육 연구는 부족한 실정이다. 이에 본 연구에서는 윤리적인 AI의 개념과 필요성을 학생들이 인지하고 프로그래밍을 통해 인공적 도덕 행위자(AMA)를 만들 수 있는 수업 프로그램을 개발하였다.

이 프로그램 안에서 학생들은 윤리적인 딜레마를 사용하여 다양한 윤리적 접근 방식에 행동이 주어졌을 때 취해야 할 행동을 토론하며 도덕성을 함양하고, 이를 프로그래밍하는 과정을 통해 프로그래밍능력 또한 높아질 수 있을 것이라 기대한다.

1-2 연구문제

본 연구는 인공지능 윤리 교육을 위해 ‘윤리적인 AI 만들기’라는 주제를 중심으로 초등 실과과와 도덕과의 내용 및 차시를 선정하여 조직하였다. 윤리적인 AI를 만드는 프로그램을 개발하기 위해 5-6학년군 도덕과 핵심가치 중 ‘공정’이라는 가치를 선정하였다. 딜레마 상황 속 토론을 통하여 학생들은 도덕적 가치에 보다 더 몰입할 수 있고, 이를 프로그래밍하는 과정에서 도덕적 가치를 내면화할 수 있다. 초등학생을 위한 인공지능 융합교육 프로그램을 개발하고 학생들에게 직접 적용하여 본 연구에서 개발한 모럴 머신 프로그램과 초등학생의 공정성 및 프로그래밍능력 간에 유의미한 결과가 있는지 알아보고자 한다.

연구문제 1. 개발한 모럴 머신 프로그램이 초등학생 공정성에 긍정적인 영향을 미치는가?

연구문제 2. 개발한 모럴 머신 프로그램이 초등학생 융합 프로젝트 프로그래밍능력 향상에 영향을 미치는가?

Ⅱ. 이론적 배경 및 선행연구

2-1 인공지능 윤리 - 윤리적인 AI

박형빈은 ‘인공지능윤리와 도덕교육(2022)’에서 초등학교의 도덕과 교육과정 내용 체계에 인공지능 윤리 교육과 관련하여 적용할 수 있는 교육을 크게 2가지로 분류한다. 첫째는 AI의 윤리, 둘째는 윤리적 AI에 대한 탐구이다. 본 연구와 관련된 분야는 ‘윤리적 AI’로 윤리적 존재로서의 인공지능의 구현을 위한 윤리 기초에 관해 탐구한다. 연계된 내용으로는 의사결정 상황에서 AI가 어떠한 도덕 판단을 내리는 것이 윤리적인가이다[3].

표 1. AMA에 관한 윤리적 접근
Table 1. Ethical approaches to AMA

Approaches	Relevant ethical theories and methods	Advantages and Limitations
Top-down	- Based on the obligatory ethics of publicism or Kant -To select a specific ethical theory and then draw out the design of algorithms and subsystems that can implement it	-Advantages: Similar to the algorithm method used by artificial intelligence to solve common problems -Limits: Difficulty in presenting specific action guidelines in reality
Bottom-up	-L. Kohlberg, A. M. Turing's virtue-ethical approach -Reinforcing through praise and reward to learn from experience	-Advantages: Advantages similar to the way humans grow into full moral actors by imitating and modifying actors through moral evaluation of specific behavioral processes. -Limits: When AI robots take a learning approach to acquire ethical tendencies through big data, the risk of becoming unethical AI robots due to distortion of given information or unethical acceptance of information
Hybrid	-Approaches to merge bottom-up and bottom-up	-Advantages: By providing pluralistic values and norms in a top-down manner and allowing AI robots to learn from various worlds and experiences, they can pursue systematic adjustments of their own. -Limits: Still unclear how to mix specifically and how to adjust to present concrete and realistic alternatives.

윤리적인 AI의 주요 개념으로는 AMA(Artificial Moral Agent, 인공적 도덕 행위자)가 있다. AMA는 스스로 도덕적 결정을 내리는 인공지능로봇이다[4]. 현재 AMA에 관한 윤리적 접근은 크게 세 가지로 표 1과 같이 구분된다.

본 연구에서는 AMA를 좀 더 직관적인 표현으로 사용하고자, Wallach, Wendell와 Allen, Colin의 저서 'Moral Machines: Teaching Robots Right from Wrong (Teaching Robots Right from Wrong)'에서[5] 사용된 'Moral Machine'이라는 용어를 사용하고자 한다. 또한, 대상이 초등학생이라는 점을 생각하여 위와 같은 복잡한 알고리즘이나 강화학습 방법을 사용하기에는 무리가 있으므로, 토론을 통해 하향식 접근으로서(Top-down) 윤리 원칙을 세우고 구체적인 이야기 상황에서 AI가 도덕적인 결정을 할 수 있도록 본인만의 모럴 머신(Moral Machine)을 프로그래밍해보는 경험을 하고자 한다.

2-3 노벨엔지니어링(Novel Engineering)

소설을 의미하는 'Novel'과 공학을 의미하는 'Engineering'의 합성어로 독서와 같이 정적인 영역과 공학이라는 동적인 영역이 합쳐진 혁신적인 융합 교육 방법이다. 노벨엔지니어링(NE)은 책 속에서 등장인물이 겪는 문제를 발견하고 이야기의 맥락에서 문제 해결을 위한 공학적 설계와 실행, 개선 및 공유의 과정을 경험하는 수업모형이다[9]. 노벨엔지니어링 수업 절차를 활용하여 학생들이 일상생활에서 있을 법한 이야기의 딜레마 상황을 통해 윤리적인 인공지능의 개발이 필요함을 이해하고 직접 설계하여 만들 기회를 가질 수 있을 것이다. Novel은 문학을 의미하고 장편 소설이나 수필도 포함되지만, 단편 동화도 포함되기에 본 연구에서는 초등학생 수준에 적합한 짧은 이야기와 소설을 활용하였다. 노벨엔지니어링의 절차는 표 2와 같다.

표 2. 노벨엔지니어링의 절차
Table 2. Procedures of Novel Engineering

Activities	The content of learning
Problem recognition	· Understanding the problem situation in the story
Scoping problems and designing solutions	· Recognizing engineering situations in problem situations and determine engineering solutions (Middle Design)
Solution testing and feedback	· Sharing mid-design deliverables and receive feedback
Improving Solutions	· Complementing the results by redesigning
Share the end result	· Making final presentation of the solution within the context of the book

2-3 인공지능-윤리 융합 교육

인공지능과 윤리 융합 교육에 대한 선행연구를 살펴보면, 사회에서 윤리적으로 기능하도록 AI를 설계하는 원리와 관련된 데이터 편향, 인공지능의 공정성, 투명성 등과 같은 주제를 중심으로 이론 교육을 연구한 논문들을 주로 살펴볼 수 있다.

하지만 본연구에서는 'AI는 사회에서 어떻게 작동해야 하는가?'와 관련하여 윤리적인 AI를 만들기 위한 프로그래밍 과정을 중심으로 한 선행연구를 찾아보았다. 아래 표 3과 같이 송수아(2022)외 2명의 논문을 찾을 수 있었다[6]-[8].

표 3. 인공지능-윤리 융합 교육 선행연구
Table 3. Previous study on artificial intelligence-ethics convergence education

Researcher	The title of the thesis	Key analysis contents	Breaking point
Song Su-ah (2022) [6]	A study on the instructional method of artificial intelligence ethics education in elementary school using self-driving coding -focusing on blended learning and definition reasoning -	If the existing artificial intelligence ethics education was focused on theoretical and discussion education, the possibility of convergence between ethics and engineering was confirmed in that the results were programmed along with theory and discussion.	Only for self-driving coding, the difficulty of generalization
Gong Eun-joo (2022) [7]	Development of AI ethics education programs reflecting social responsibility	Development of educational programs that allow coding learners to learn to consider the social responsibility and ethical implications of AI development together in the process of writing the programs they create.	High school-level programs, Not applicable to actual training sites
Kim Joo-hyun, Moon Sung Hwan (2021) [8]	Artificial intelligence education program for elementary school students using Novel Engineering	For 4th grade elementary school students, Novel Energizing is used to learn the main functions of artificial intelligence by unplugging. Finding easy and fun algorithm learning possibilities suitable for elementary school level	Lack of generalization as an artificial intelligence leading school, the 4th grade of elementary school

Ⅲ. 모럴 머신 교육 프로그램 연구방법

3-1 연구대상

본 연구는 초등학생 도덕성 개발에 영향을 미치는 모럴 머신 프로그램을 개발하고 프로그램을 적용한 뒤에 학생들의 도덕성과 인공지능 프로그래밍 능력에 미치는 효과를 분석하고자 한다. 이를 위해 A지역 초등학생 중 인공지능과 SW(Software) 교육에 관심있는 5, 6학년 20명의 학습자를 대상으로 프로그램을 적용하였고 연구 대상의 구성은 표 4와 같다.

표 4. 연구 대상 구성
Table 4. Composition of study participant

Grade	Male	Female	N
5th	11	0	11
6th	6	3	9
N	17	3	20

3-2 학습자 실태 분석

프로그램을 적용하기 전에 학습자들의 인공지능과 인공지능 윤리에 대한 인식을 알아보기 위해 사전 설문조사를 실시하였고, 실태를 분석한 결과는 표 5와 같다.

표 5. 인공지능 교육에 대한 흥미
Table 5. Interest in artificial intelligence education

	Very high	High	General	Low	Very low	Total
AI education	13 (65%)	6 (30%)	1 (5%)	·	·	20 (100%)

인공지능 교육에 대한 흥미도는 65%의 학생이 매우 높음에 응답하였고, 30%의 학생이 높음에 응답하였다. 인공지능에 관심이 많은 학생들이고, 자발적으로 수업에 참여했기 때문에 대부분의 학생이 인공지능 교육에 흥미가 높았다. 표 6은 프로그래밍 도구 사용 경험에 대한 설문결과이다.

표 6. 프로그래밍 도구 경험
Table 6. Experience with programming tools

	Block coding	App in venter	Physical Computing	Text coding
N	20(100%)	4(20%)	13(65%)	9(45%)

프로그래밍 도구 경험에 대한 설문은 중복 투표를 할 수 있도록 설정하였기에 각 항목의 확률은 20명을 전체로 보고 구한 값이다. 위 표와 같이 프로그래밍 도구 경험에서 모든 학생이 블록코딩(엔트리, 스크래치) 경험이 있다고 응답하였다. 피지컬 컴퓨팅 경험이 있는 학생도 65%를 차지하였고 텍스트 코딩도 45%로 절반 가까운 학생들이 텍스트 코딩 경험이 있다고 응답하였다. 다음은 '인공지능 윤리'에 대한 인식을 워드클라우드로 정리한 결과이다. 인공지능에 대한 흥미가 높고 프로그래밍 도구 사용 경험이 많은 반면 학생들은 인공지능 윤리라는 개념도 생소함을 느끼고 있었고 관련된 단어를 떠올리는 데 어려움이 있었다. 많은 학생이 인공지능 윤리를 들었을 때 떠오르는 단어가 없다고 답했으며, 그 외에 인간의 안전, 법칙, 악용이라는 단어를 떠올리고 있었다. 이에 본 연구에서는 학생들이 쉽게 마주할 수 있는 딜레마 상황을 해결해보고 본인이 직접 모럴 머신을 프로그래밍하는 경험을 통해 인공지능 윤리의 개념을 바르게 이해하고, 윤리적인 AI의 필요성을 느끼게 하고자 한다.

*The students answered in Korean.
그림 1. 인공지능 윤리에 대한 인식 단어 모음
Fig. 1. A collection of words to recognize AI ethics

3-3 평가 설계

본 교육프로그램이 초등학생의 도덕성에 어떤 영향을 미치는지를 확인하기 위하여 양적 평가와 질적 평가를 통해 효과성을 알아보고자 한다. 양적 평가로는 공정과 관련된 도덕성 평가지를 교육프로그램 사전 사후 검사하여 도덕성의 변화 여부를 분석한다. 질적 평가로는 수업 중 학생들이 프로그래밍한 두 가지 모럴 머신을 평가하여 학습자의 프로그래밍 능력이 향상되었는지 평가한다. 표 7은 교육 프로그램의 평가 설계를 표로 정리한 내용이다.

표 7. 평가 설계
Table 7. Evaluation design

Evaluation method	Type	Purpose	Time
A quantitative evaluation	Moral test [10]	Analysis of morality changes after application of training program.	Before and after the program starts.
A qualitative evaluation	Performance evaluation of making your own moral machine.	Evaluate learners' moral judgment and programming ability by evaluating their own morphological machine-making output.	After the novel engineering utilization class.

Ⅳ. 모럴 머신 교육 프로그램 연구결과

4-1 관련 교과 및 단원

6학년 도덕의 핵심 가치는 '공정'과 '책임'이다. 본 연구에서는 도덕적 딜레마 상황에서 선택의 문제와 연결될 수 있는 '공정'이라는 가치에 초점을 맞춰 교육프로그램을 구성하고자 한다. 또한 6학년 실과의 소프트웨어 교육 17차시 중 '실생활 문제 해결 프로그램 설계'와 연계하여 이전에 배웠던 프로그래밍 능력(순차, 선택, 반복 구조)뿐만 아니라 인공지능과 관련된(예측하기-회귀분석) 기능을 학습시키고자 한다.

표 8. 관련 교과 내용 분석
Table 8. Analysis of related subject content

	What's inside	
Session	14 sessions (8 sessions for morality, 6 sessions for practical course)	
Related subject	Morality	Practical course
Area (Units)	The relationship between society and the community (4. A Fair Life) + Competency-focused, autonomous integration unit	Technical system (4. Software for communication)
Achievement Standards	[6 morality 03-01] Understand the meaning of fairness and the necessity of a fair society, and develop a willingness to practice living fairly in everyday life.	[6 practical course 04-11] Understand the structure of sequential, selection, repetition, etc. in the process of creating a program to solve problems.

4-2 차시별 수업 내용

본 연구의 구체적인 목표는 다음과 같다. 첫째, 본 연구에서 개발한 프로그램은 초등학생의 공정성을 함양시킬 수 있다. 둘째, 본 연구에서 개발한 프로그램은 학생들의 인공지능 프로그래밍능력을 향상시킬 수 있다. 프로그램은 표 9와 같다.

표 9. 수업 절차
Table 9. Instructional procedures

Learning content by lesson			
session	Model		Learning content
1	Conceptual Analysis Class Model		· Understanding the Meaning of Fairness
2	a value-deepening curriculum		· Learn about the moral subjects and objects of fairness (AMA education) and the importance of fairness (5 principles)
3-4	Novel engineering	Problem recognition	· Discover the problem through the 'Running Representative' story · Group discussions to design a fair way to solve problems (attribute, weight extraction)
		Scoping problems and designing solutions	
5-6		Solution testing and feedback	· Making your own Moral Machine1 (Rule-based coding using algorithms)
7-8		Improving solutions & Share the end result	· Modifying your own Moral Machine (Use the Predicting Model (Regression Analysis)
9-10	Novel engineering	Problem recognition	· Discovering Problems through 'Rich Old Man of Unmanned Island' · Group discussions to design a way to solve the problem situation of a story (dilemma situation) (attributes, weight extraction)
		Scoping problems and designing solutions	
11-12		Solution testing and feedback	· Make Your Own Molecular Machine2 (Use Self-Prediction Model (Regression Analysis))
13-14		Share the end result	· Presenting and evaluating the final results - Look at the group's moral machine and see if the difference in attribute values and weights made the results different - To determine if the other group's moral machine is fair

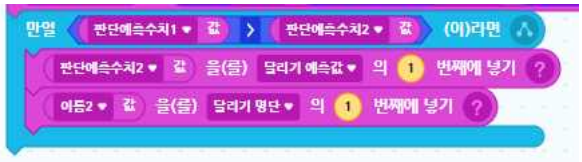
4-3 주요 프로그래밍(엔트리) 코드

본 연구에서 쓰이는 프로그래밍은 크게 2가지로 규칙기반 프로그래밍과 인공지능 예측하기 모델(회귀분석)을 활용한 프로그래밍이다. 본 수업 3-4차시 중 학생들은 '달리기 대표'라는 이야기를 읽고 반대표로 달리기 시합에 나가는 친구를 뽑는 과정에서 딜레마 상황을 겪게 되며 도덕적 판단을 위한 토론을 통해 속성값과 가중치를 결정하게 된다. 이 속성값과 가중치를 통하여 7-8차시 수업에서 도덕적으로 판단하는 AI를 만들게 된다. 아래 그림 2~그림 4는 본 수업 7-8차시에서 활용되는 인공지능 예측하기 모델을 활용한 프로그래밍 예시 코드이다.

* The entry program is a block coding program that supports Korean.

그림 2. 변숫값 저장 코드
Fig. 2. Saving variable values code

위 그림 2의 주요 코드는 반 대표로 달리기에 출전할 선수를 공정한 방법으로 뽑기 위해 학생들이 토론을 통해 뽑은 속성값(학년, 빠르기, 실제 등수)을 변수로 받아서 저장하는 과정이다. 속성값을 모두 받았으면, 예측하기 모델(회귀분석)을 통한 식을 활용하여 판단예측 수치에 결과값을 저장하게 된다.



* The entry program is a block coding program that supports Korean.

그림 3. 판단예측 수치 비교 코드
Fig. 3. Comparison of judgment prediction figures code

위 그림 3의 주요 코드는 판단예측 수치로 저장된 학생1과 학생2의 결과값을 비교하여 더 결과값이 낮은(순위가 높은) 학생의 이름을 달리기 명단(리스트)에 넣는 과정이다.

* The entry program is a block coding program that supports Korean.

그림 4. 공정한 달리기 대표 판단 코드

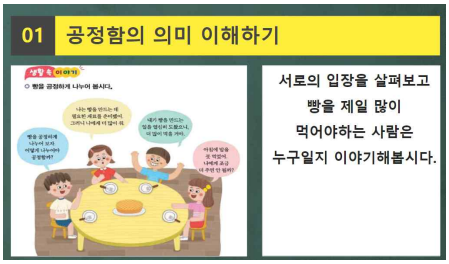
Fig. 4. To judge a fair running representative code

위 그림 4의 주요 코드는 사용자가 판단하기 버튼을 눌렀을 때, 달리기 명단(리스트)에 첫 번째 기록된 학생의 이름을 인공지능이 말하게 하는 과정이다.

4-4 프로그램 적용

1) 공정함의 중요성 알기

1차시는 공정함의 의미를 이해하는 차시로 빵을 공정하게 나누는 방법을 생각하며 서로의 입장에서 빵을 제일 많이 먹어야 하는 사람은 누구인지 결정하는 과정을 통해 학생들은 '공평'과는 다른 '공정'의 가치를 이해했다. 2차시에서는 현재 인공지능 사용 결과 공정하지 못했던 사례를 살펴보며 '인공지능 윤리'와 '공정'을 연결시켜 생각할 수 있도록 돕고 이를 통해 인공지능도 도덕적 주체로서의 지위를 갖게 될 필요성이 있음을 인지하게 하였다. 마지막 활동으로는 공정함의 원리를 알아보며 공정함의 중요성을 내면화해 보았다. 공정함의 중요성 교육자료는 그림 5와 같다.



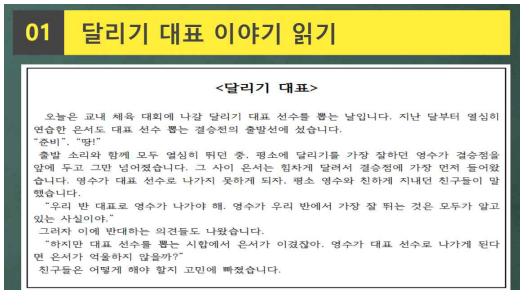
* It's an educational material produced in Korean.

그림 5. 공정함의 중요성 교육 장면

Fig. 5. The importance of fairness educational scene

2) '달리기 대표' 이야기를 통한 모럴 머신 만들기

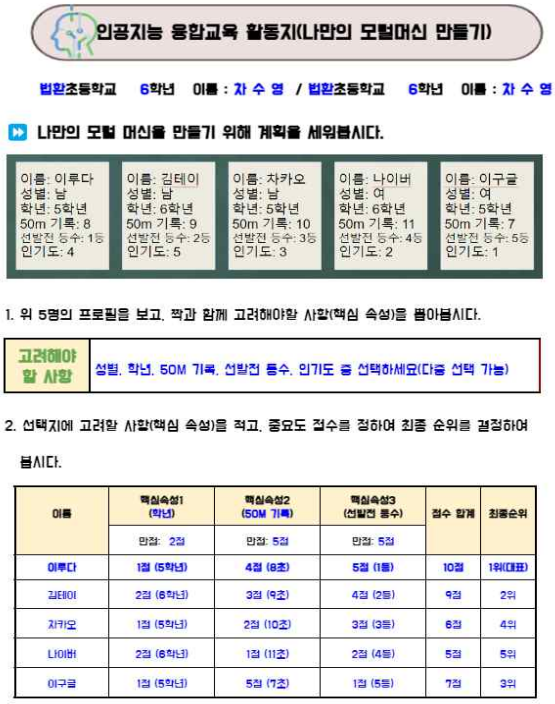
첫 번째 프로젝트 활동 '달리기 대표' 이야기를 통한 모럴 머신을 프로그래밍하는 활동이다. 그림 6과 같이 3-4차시에서는 6학년 도덕 교과 4. 공정한 생활에 제시된 이야기 '달리기 대표'를 읽고 이야기 속 도덕적 딜레마 상황을 통해 달리기 대표를 공정하게 뽑으려면 어떤 요인들이 필요한지 고려하였다. 고려해야 할 요인을 선정하기 위해 7가지 속성이 담긴 5명의 프로필을 그림 7과 같이 제공하였다. 노벨엔지니어링 수업에서 "내가 생각한 문제를 나와 동료와 같이 해결"[11] 할 때, 학생들의 마음의 변화를 더 잘 끌어낼 수 있으므로 학습형태를 모둠으로 하여 모둠 안에서 학생들은 친구들과 토론하며 '공정한 달리기 대표' 선정에 필요한 속성을 선택하고 속성별 가중치를 얼마나 배분할지 의논하였다.



* It's an educational material produced in Korean.

그림 6. '달리기 대표' 이야기를 통한 문제 발견 장면

Fig. 6. Scene of finding problems through the story of "Running Representative"



* It's an educational material produced in Korean.

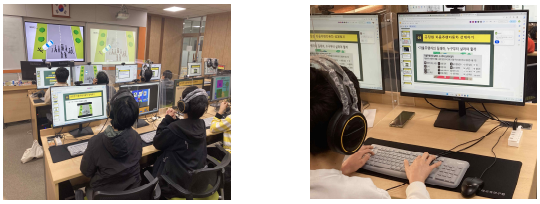
그림 7. 달리기 대표 모럴 머신 제작 장면

Fig. 7. The scene of making a running representative moral machine.

5-6차시에서는 친구와 논의한 속성과 가중치를 바탕으로 학생들이 스스로 프로그래밍해보게 하였다. 학생들은 엔트리의 인공지능 모델 학습 기능을 사용해 본 경험이 없기 때문에 단순 규칙기반 알고리즘을 활용하여 프로그래밍하였다. 이후 7-8차시에서 인공지능 모델 학습하기- 예측하기(회귀분석)를 도입하고 주어진 순서를 통해 엔트리의 예측하기 모델을 활용하여 '공정한 달리기 대표'를 뽑는 모럴 머신을 프로그래밍해보았다. 프로그래밍 후 발표하는 시간을 가지며 도덕적 판단을 위해 필요한 속성과 왜 이 속성에 더 가중치를 높게 주었는지를 공유하는 시간을 가졌다. 또 모둠별 피드백을 통해 우리 모둠에서 프로그래밍한 결과가 공정한지 생각해보는 시간을 가졌다.

3) '무인도 부자노인' 이야기를 통한 모럴 머신 만들기

두 번째 프로젝트 활동은 '무인도 부자노인' 이야기를 통한 모럴 머신을 프로그래밍하는 활동이다. 우선 9-10차시에서는 동기유발을 위해 본 수업의 이야기와 비슷한 맥락을 가진 '트롤리 딜레마' 상황 제시를 통해 그림 8과 같이 내가 자율주행 자동차를 설계한다면 어떤 사람을 먼저 살릴지 생각해보고 의견을 나누는 시간을 가졌다. 학생들은 나이, 인종, 질병 유무 등의 속성을 추출하였고 여러 속성 중 어떤 속성에 가중치를 두어야 할지 간단하게 발표하였다.

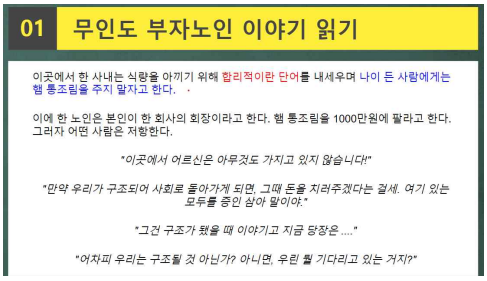


* This is a scene from a class conducted in Korean.

그림 8. 트롤리 딜레마 도덕적 딜레마 교육 장면

Fig. 8. Trolley dilemma moral dilemma educational scene

그리고 그림 9의 '무인도 부자노인' 이야기를 읽으며 무인도에서 식량을 배분해야 하는 도덕적 딜레마 상황을 제시하였다. 학생들은 모둠 활동을 통해 어떻게 식량을 공정하게 배분해야 할지 다양한 의견을 공유하였다. 나이, 성별에 관계없이 똑같이 식량을 배분해야된다는 의견부터 경제력이나 건강상태, 몸무게 등에 따라 차등을 두어 식량을 배분해야 한다는 의견이 나왔다.



* It's an educational material produced in Korean.

그림 9. '무인도 부자노인' 이야기[12]

Fig. 9. The story of "The Unmanned Old Man."

11-12차시에서는 전 차시에서 의견을 공유한 내용을 바탕으로 식량을 공정하게 배분하는 모럴 머신을 프로그래밍하였다. 프로그래밍 후, 13-14차시에서 학생들은 나만의 모럴 머신을 발표하며 친구들과 생각 공유를 통해 식량을 공정하게 배분하기 위해 다른 모둠은 어떤 속성으로 예측하기 모델을 제작하였는지 알 수 있었고 '다른 모둠이 프로그래밍한 모럴 머신은 공정한가?'에 대해 평가해보고 공정하다면 왜 공정하다고 생각하는지, 그렇지 않다면 왜 그런지 구체적인 피드백을 주고받았다. 이를 통해 다시 한번 본인이 생각하는 '공정'의 가치와 개념을 명확히 하고 내면화하는 시간을 가졌다.

4-5 모럴 머신 프로그램 적용 결과 분석

1) 공정에 대한 도덕성 변화 분석

모럴 머신 프로그램 적용 후 학생들의 공정에 대한 인식을 7개의 문항으로 변화 여부를 분석한 결과는 표 10과 같다. 교육 전 공정에 대한 인식 평균값은 24.35에서 교육 후 27.25로 2.9점 증가하였다. 또한 p값도 0.01보다 작아 모럴 머신 프로그램 적용 후 공정에 대한 인식변화에 유의미한 결과가 있음을 확인하였다. 특히 '내가 가지고 있는 것을 원하는 사람들에게 공평하게 나누어준다.' 문항에서 14.1%, '문제를 평화롭고 정당하게 해결하려고 애쓴다.' 문항에서 14%, '친구들이 싸울 때 양쪽 얘기를 모두 들어 본다.' 문항에서 19%, '서로가 하고 싶은 일을 차례로 정해 교대로 한다.' 문항에서 14%가 높아졌다. 모럴 머신 프로그램을 통해 학생들의 공정에 대한 인식이 매우 높아졌음을 알 수 있다.

표 10. 도덕성 공정의 인식과 관련된 설문에 대한 T 검정 분석 결과

Table 10. Results of the T-test analysis for the questionnaire related to the perception of the moral process

(N=20)

	Time	M	SD	t(p)
the perception of moral fairness	pre	24.35	4.727	0.002(<.01)**
	post	27.25	4.387	

**p<.01

2) 모럴 머신 프로그래밍 프로젝트 수행평가 분석

본 교육프로그램에서 학생들은 두 가지의 도덕적 딜레마 상황을 당면하였고 직접 프로그래밍 제작을 통해 문제를 해결하였다. 우선 첫 번째 프로젝트에서는 '달리기 대표' 이야기를 읽고 달리기 대표를 선출하기 위한 속성을 추출하고 가중치를 배분하였고 모럴 머신을 프로그래밍하기 위한 구상 활동과 더불어 단순 규칙 기반 알고리즘을 통해 프로그래밍하였다. 두 번째 프로젝트에서는 '무인도 부자노인' 이야기를 읽고 공정하게 식량을 나누기 위한 속성을 추출하고 첫 번째 프로젝트 후반에 다뤄진 인공지능 모델 - 예측하기(회귀분석)을 활용하여 주어진 값이 달라져도 공정하게 식량을 배분하는 모럴 머신을 프로그래밍하였다. 두 프로젝트에서 제작된 모럴 머신을 문제해결력, 인공지능 기술 적용, 창의성, 도덕적 판단력, 완성도 총 5가지 영역으로 나누어 수행평가를 진행하였다. 모든 하위 변인에서 점수가 증가하였고 특히 인공지능 기술 적용에서는 첫 번째 프로그래밍 후 인공지능 모델-예측하기(회귀분석)을 도입하였기에 '인공지능 기술 적용' 항목에서 점수 증가폭이 매우 높았다. 본 프로그램을 통해 학생들의 전체적인 프로그래밍능력이 향상되었고 자세한 분석 결과는 표 11과 같다.

표 11. 모럴 머신 프로그래밍 프로젝트 수행평가 분석 결과
Table 11. Results of performance evaluation of the moral machine programming project

(N=20)

Subvariable (Out of 20 points each)	The Moral Machine of the 'Running Representative'	The Moral Machine of the 'Rich Old Man of Unmanned Island'	Increase
Problem-solving ability	13.5	15	1.5
Application of AI	0	14.5	14.5
Creativity	10	14	4
Moral judgment	10	13	3
Degree of completion	12	14.5	2.5

V. 결론 및 제언

본 연구에서는 초등학생이 인공지능 윤리에 대한 이해를 바탕으로 윤리적인 AI를 프로그래밍하면서 공정성이라는 도덕적 가치에 대해 깊이 탐구하고 내면화할 수 있는 프로그램을 설계하였다. 이를 위해 6학년 도덕과 실과의 내용 체계를 분석한 후, 인공지능과 융합하여 도덕성을 함양시키고 프로그래밍 능력을 키울 수 있는 14차시의 프로그램을 개발하였다.

A지역 초등학생 중 인공지능과 SW 교육에 관심 있는 5,6학년 20명의 학습자를 대상으로 14차시의 프로그램을 2차시씩 7주간 적용하였고 학습자들은 공정과 인공지능 윤리를 연결시키며 나만의 모럴 머신1, 2를 프로그래밍하였다. 수업 전후로 도덕성 평가지를 활용하여 질적평가와 프로젝트 수행평가를 통한 양적평가 결과 수업 후 '공정'과 관련된 도덕성 및 프로그래밍능력이 향상되었음을 알 수 있었다.

본 연구는 단순한 인공지능 윤리의 이론 교육 방향이 아닌 인공지능과 윤리교육을 결합하여 실제 프로그래밍해보는 과정과 방법을 제안했다는 점에서 의의가 있다. 도덕적 딜레마 상황에서 학생들이 직접 윤리적인 인공지능을 설계하는 과정에서 학생들은 '공정'이라는 도덕성을 깊게 이해하게 되었고 이를 직접 인공지능으로 구현하면서 프로그래밍능력이 높아지는 것을 보아 초등학교 도덕 수업뿐만 아니라 인공지능 수업까지 일선 교육 현장에 의미 있게 활용될 수 있을 것으로 기대한다.

다만 관련 연구로서 본 연구에서 제시된 프로그램을 바탕으로 다른 가치도 적용될 수 있는지 연구할 필요가 있다. 본 연구는 많은 도덕 가치 중 '공정'이라는 가치에만 국한되어 프로그램이 제작되고 평가했기 때문이다. 더 나아가 도덕 교과에서 제시된 가치를 중심으로 인공지능 윤리를 연결하여 가르칠 수 있는 일반화된 프로그램 및 모형에 대한 연구가 요구된다.

감사의 글

이 논문은 2023학년도 제주대학교 교원성과지원사업에 의하여 연구되었음.

참고문헌

[1] Ministry of Education. Elementary and Middle School + Curriculum + General Theory + Draft [Internet]. Available: https://moe.go.kr/boardCnts/viewRenew.do?boardID=141&boardSeq=93001&lev=0&searchType=null&statusYN=W&page=1&s=moe&m=040401&opType=N.

[2] Ministry of Education. Elementary School Curriculum Draft [Internet]. Available: https://moe.go.kr/boardCnts/viewRenew.do?boardID=141&boardSeq=93001&lev=0&searchType=null&statusYN=W&page=1&s=moe&m=040401&opType=N.

[3] H. B. Park, *AI Ethics and Moral Education*, Seoul: CIR, 2022.

[4] H. C. Choi, S. Y. Byun, and H. J. Shin, "The Ethical Principles for the Development of Artificial Moral Agent - Focusing on the Top-down Approach -," *Journal of Ethics*, No. 111, pp. 31-53, December 2016. https://doi.org/10.15801/JE.1.111.201612.31

[5] B. S. Kim, The Effect of Novel Engineering on Reading skills, Problem Solving and Engineering Creativity, Master's Thesis, Korea National University of Education, Cheongju, February 2020.

[6] W. Wallach and C. Allen, *Moral Machines: Teaching Robots Right from Wrong*, New York, NY: Oxford University Press, 2009.

[7] S. A. Song, A Study on Teaching Method of AI Ethics Education Applied to a Self-Driving Coding in Elementary School -With Emphasis on Blended Learning and Moral Reasoning of Justice-, Master's Thesis, Gyeongsang National University, Jinju, August 2022.

[8] J. Kim and S. Moon, "Development of an AI Education Program Based on Novel Engineering for Elementary School Students," *The Journal of Korea Elementary Education*, Vol. 32, No. 1, pp. 425-440, March 2021. http://dx.doi.org/10.20972/kjee.32.1.202103.425

[9] E. J. Kong, Development of Artificial Intelligence Ethics Education Program Reflecting Social Responsibility, Master's Thesis, Korea National University of Education, Cheongju, February 2022.

[10] J.-J. Jeong, "The Development and Validation of the Moral Intelligence Test for Elementary School Children," *The*

Journal of Elementary Education, Vol. 22, No. 1, pp. 249-269, February 2009.

[11] K. C. Hong, “Novel Engineering as a Convergence Class Model,” in *Proceedings of 2019 Korean Association for Qualitative Inquiry Fall Conference*, Jeonju, pp. 52-56, October 2019.

[12] D. S. Kim, *A Gray Person*, Seoul: Yoda, 2017.

문수지(Moon Su Ji)

2013년 : 제주대학교 초등체육교육과 (학사)

2023년～현 재: 새서귀초등학교 교사

2023년 9월～현 재: 제주대학교 교육대학원 인공지능융합교육전공(석사과정)

※관심분야 : 인공지능융합교육, SW·AI 교육

김성백(Seong Baeg Kim)

1989년 : 서울대학교 컴퓨터공학과 (공학사)

1991년 : 서울대학교 컴퓨터공학과 (공학석사)

1995년 : 서울대학교 컴퓨터공학과 (공학박사)

1996년～현 재: 제주대학교 컴퓨터교육과 교수

2020년～현 재: 제주대학교 교육대학원 인공지능융합교육전공 교수

※관심분야 : 컴퓨팅시스템, SW·AI 교육, 컴퓨터교육 등