JUNYEOB BAEK

Ph.D. Student @ KAIST

✉ wnsdlqjtm@kaist.ac.kr 🔗 Junyeob ⌨ dion-jy 🌐 dion-jy.github.io

Last updated: Oct. 2025

RESEARCH INTEREST

My research interests lie in building **lifelong developmental agents** that acquire emergent cognitive capabilities through self-explored experiences, while advancing interpretability for cognitive monitoring and control, and alignment techniques. To achieve this vision, I focus on three core research questions:

- *Q1. How can we build interpretable systems to monitor and understand AI's internal cognitive processes?*
  Keywords: **World Models** & **Interpretability**

- *Q2. How can AI develop unique intelligence through self-explored experience rather than supervision?*
  Keywords: **Open-ended Learning** & **Intrinsic Motivation**

- *Q3. How can we ensure AI shares human values and develops in directions that benefit humanity?*
  Keywords: **AI Safety** & **Alignment**

EDUCATION

Korea Advanced Institute of Science and Technology(KAIST) **Mar. 2025 – Present**
Ph.D. in Computer Science - ***Advisor:*** *Sungjin Ahn* *Daejeon, Korea*

Korea Advanced Institute of Science and Technology(KAIST) **Mar. 2023 – Feb. 2025**
M.S. in Artificial Intelligence - ***Advisor:*** *Sungjin Ahn* *Daejeon, Korea*
- **Thesis**: Dreamweaver: Learning Compositional World Models from Pixels

Hanyang University(ERICA) **Mar. 2014 – Feb. 2018**
B.S. in Robot Engineering - ***GPA*** *- 4.29 / 4.5* *Ansan, Korea*
- *Summa Cum Laude*
- *Capstone Project1: Disaster and Rescue Robot Simulation Modeling(Darrsm) - 2nd Prize in College*
- *Capstone Project2: Walk Yourself, Toddler!(WalkYTo) - 3rd Prize in Department*

RESEARCH EXPERIENCE

Visiting Researcher **Jun. 2025 – Aug. 2025**
Mila-Quebec AI Institute *Montreal, Canada*
- Collaborating with *Yoshua Bengio*'s research group and *LawZero* team members
- Conducting research project on **Mechanistic Interpretability and Chain-of-Thought Faithfulness in LLMs**
- **Main Collaborators:** *Joumana Ghosn*, *Pierre-Luc St-Charles*, *Marc-Antoine Rondeau*, *Minsu Kim*

PUBLICATIONS

1. **Junyeob Baek**, Hosung Lee, Christopher Hoang, Mengye Ren, and Sungjin Ahn. "Discrete JEPA: Learning Discrete Token Representations without Reconstruction". In: *International Conference on Machine Learning (ICML) Tokenization Workshop* (2025).

2. Donghwan Chi, Hyomin Kim, Yoojin Oh, Yongjin Kim, Donghoon Lee, Daejin Jo, Jongmin Kim, **Junyeob Baek**, Sungjin Ahn, and Sungwoong Kim. "Slot-MLLM: Object-Centric Visual Tokenization for Multimodal LLM". *Preprint, Under Review* (2025).

3. **Junyeob Baek**, Yi-Fu Wu, Gautam Singh, and Sungjin Ahn. "Dreamweaver: Learning Compositional World Models from Pixels". In: *International Conference on Learning Representation (ICLR)* (2025).

4. Chang Chen, **Junyeob Baek**, Fei Deng, Kenji Kawaguchi, Caglar Gulcehre, and Sungjin Ahn. "PlanDQ: Hierarchical Plan Orchestration via D-Conductor and Q-Performer". In: *International Conference on Machine Learning (ICML)* (2024).

THESIS

- **Junyeob Baek**. "Dreamweaver: Learning Compositional World Models from Pixels". Master's Thesis, KAIST, Advisor: Sungjin Ahn (Feb. 2025).

ACADEMIC SERVICE

Conference Reviewer

- International Conference on Learning Representations (ICLR) 2025

Workshop Organizer

- 2nd KAIST-Mila Prefrontal AI Workshop (6 invited speakers incl. Yoshua Bengio), Link

TEACHING EXPERIENCE

Teaching Assistant

CS377: Introduction to Reinforcement Learning, KAIST *Spring 2025*
- Outstanding Teaching Assistant Award recipient

PROFESSIONAL EXPERIENCE

MakinaRocks **Jun. 2021 – Jun. 2022**
ML Research Engineer(Robotics Focused) *Seoul, Korea*
- **Optimization of Welding Task and Motion Planning in Vehicle Manufacturing Process**
  * Challenges: 1. Large-scale and high complexity in Optimization Problem. 2. Required high-accuracy planning in highly cluttered environments. 3. Required tight Time Window for control
  * [*Contribution 1*] GA-based Multi-objective Optimization Module for Welding-point Assignment
  * [*Contribution 2*] Multi-Robot Interlock Scheduling System with Swept-volume-based collision detection

Syscon, Co., Ltd. **Dec. 2017 – Jun. 2021**
Robotics Software Engineer/Research Strategy Manager *Incheon, Korea*
- **AMR Navigation System Design & Product Development**
- [*Field Exp.*] **Successful Deployment for Smart Factory Project**
  * More than 300 delivery tasks per day with 14 robots
- [*Research Exp.*] **Learning-based High-accuracy Object Pose Estimator(Dock Pose Estimation)**
- [*Research Exp.*] **Distilled Meta Policy Navigation System for Multi-behavior Skillset and its Strategies**