Introduction of Simultaneous Recognition and Assessing (SRAA)

Yunfan ‘Damien’ Wang

Research Associate at Neo Scholar and UNSW CSE, Mechatronics & Computer Science student at UNSW

[wyfwhplh@126.com](mailto:wyfwhplh@126.com)

This is an introduction to a new software architecture called simultaneous recognition and assessing.

1. **Background knowledge and problem formulation**

With the AI taking up the top words in google search, an era of intelligent system powered by means machine training is, no longer around the corner, out in the air. However, according to Allen Collin, an AI with no moral constraint or sense of ethics can constitute intrinsic doubt and will never find its niche in human society [1]. So, it comes as no surprise when there has been a growing concern about intelligent machines’ reliability. According to a survey report by Khari Johnson, people are feeling discontent with the fast pace at which the AI industry is upgrading [2]. Such concerns evolve into a gigantic hindrance to the implementation of many AI-related products. According to Johannes, a member of the ethics fellow in Stanford University, the before many extreme challenges can be answered, it is not reliable to take self-driving cars.[3]

There are also ethics problem that’s challenging existing social values. According to Jack Boeglin [4], autonomous cars equipped with camera and recognition system will impair peoples’ basic rights such as freedom and privacy.

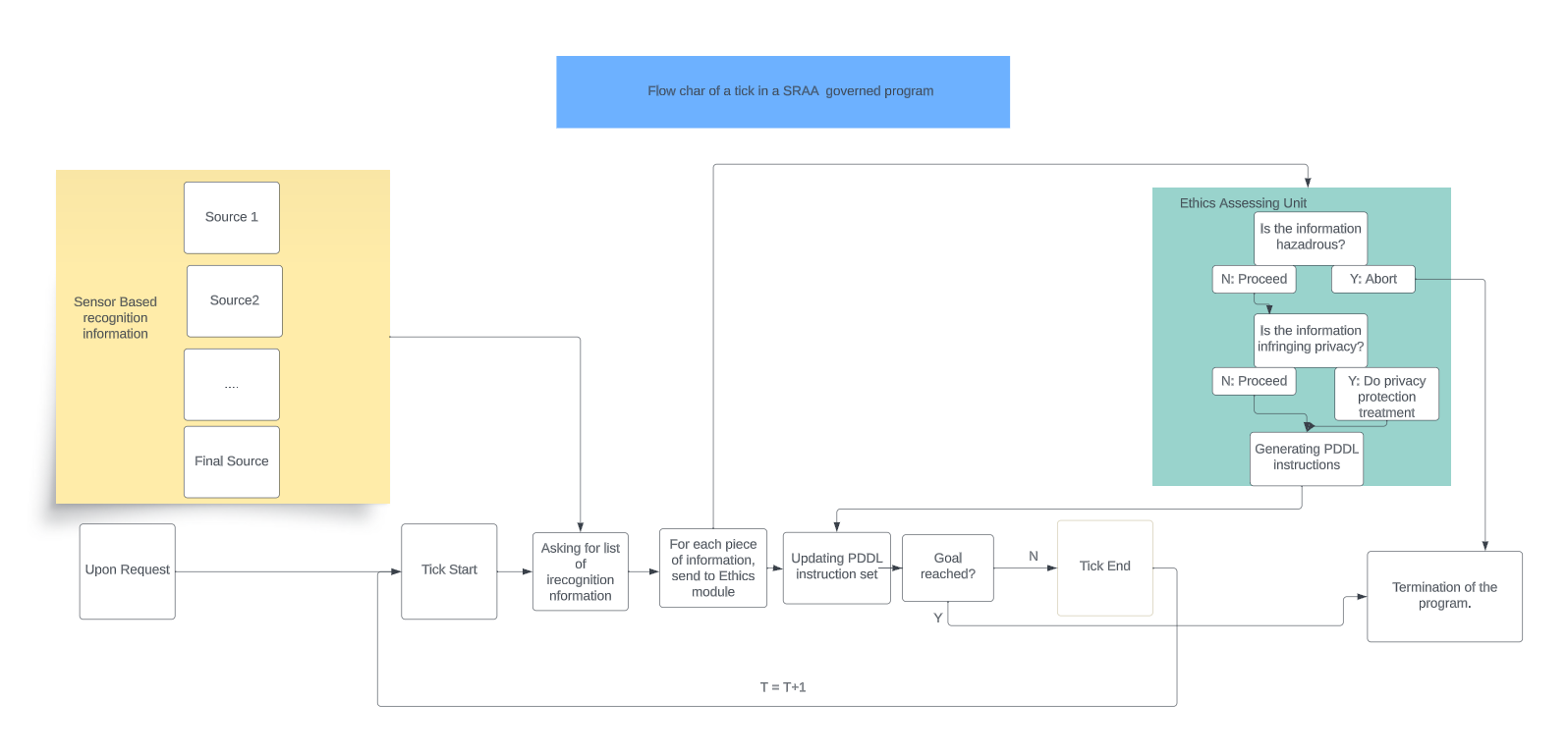
Apart from safety concerns, industry needs and moral concerns, machine ethics remains a key hinge in the desired birth of AGI (Artificial General Intelligence) [5]. According to Shahana, machine ethics come into a strong play when talking about AGI in terms of responsibility, privacy, and decision-making. In the end of that paper, Author comes into a strong conclusion: “As the AGI landscape evolves, ethical considerations must remain at the forefront of every decision, guiding the responsible development and deployment of this groundbreaking technology.” That conclusion can well encapsulate the problem we, as ethics researchers, are putting efforts to resolve. A trustworthy AI with ethics can be most decisive and thus most desirable.

1. **Simultaneous Recognition and Assessing**

2.1 Source of inspiration and brief

In 2023, I attended VIP (Vertical Integrated Program) led by Prof. Maurice Pagnucco and A.Prof Yang Song, which aims to explore a wide range of aspects in machine ethics and the use of sensor information to assist a more humane decision making AI. Inspired by the final product whose software is in GitHub, I put forward a software architecture called simultaneous recognition and assessing (SRAA). It is because although much software has a wide range of modules, but never has an architecture addressing the importance of ethical module been summarized. In other words, never has an architecture been brought forward addressing the incorporation of ethical decision module and information processing unit.

* 1. Architecture brief

Figure 2.1 expresses the idea of simultaneously incorporating recognition and ethics assessing module.

*Figure 2.1 The flow chart of SRAA architecture*

Upon requesting, a program is scheduled to continuously receive sources of information and PDDL instructions (Planning Domain Definition Language), then achieve some goal. In each tick of the program, it will first ask sensors to update the source of information, then it sends the updated information list to the Ethics assessing unit. Engineers can insert different ethics decision blocks in the ethics assessment unit. After the decisions are made, it will either update the PDDL instructions or exit the ticking loop. Then the program will keep on deciding whether to start another loop base on whether goal condition is reached.

* 1. Yolo v7 implementation.[7]

In the GitHub repository, under “src/SRCC” is my implementation of YOLO v7 based PDDL planning system after incorporating YOLO v7 with PDDL instruction generators. It composes of a nursery house scenario PDDL generation PDDL engine and a yolo based assessing module. Some of the function remains to be done, but the skeleton is there to show. Note, this GitRepo does not contain a training module, which could be useful to boost the Ethics Assessing Unit, which will be discussed in detail in the following “reflection and upgrades” section.

**3. Reflection and upgrades**

In general, I am very satisfied with this set-up, as it is very much an industry biased solution. Note that in Figure 2.1, the status is not generated, but rather, updated. This significantly boosts the error-tolerance. If one sensor goes wrong, or has discontinuation in data flow, the whole system will not shut down. Also using method of updating is beneficial for speed boosting, we can use the token to check if sensor information needs to be updated, rather than constantly generating repeating information. A similar idea of updating data is also present in ethics assessing module.

There is also clear room for upgrading. For the Ethics assessing part, there are clear patterns for data input (segmented video frame or extracted token by language model), we can use machine learning method to frame and train the whole ethics assessing unit. Multimodality fusion measures can be taken here to fuse multiple source of information to a output of PDDL instruction[6]. Which, instead of going through a series of if-statements, enables an end-2-end direct output. However, some critical hazardous condition must be checked first to proceed into the remainder of the assessing unit.

**4. Conclusion**

The exploration and development of the Simultaneous Recognition and Assessing Architecture (SRAA) underscores the vital role of ethics in the realm of Artificial Intelligence, particularly as we edge closer to the realization of Artificial General Intelligence (AGI). This paper, drawing inspiration from the Vertical Integrated Program and incorporating elements like YOLO v7 into a PDDL-based planning system, highlights the intricate balance between technological advancement and ethical considerations. Our journey through the implementation of SRAA in a nursery house scenario has reinforced the notion that ethical considerations are not just supplementary but are central to the development of AI systems. Looking forward, the potential integration of machine learning methods to train the Ethics Assessing Unit promises a leap towards more sophisticated, nuanced decision-making processes.

In conclusion, the development of SRAA could represent a promising stride in embedding ethical decision-making within AI systems. It's a testament to the necessity of intertwining ethical considerations with technological advancements. Further work shall be done to ensure the landing of this yet-theoretical architecture.

**5. GitHub Repository**

https://github.com/Yunfan-Wang/UNSW-VIP-ATHOME-CVME

**6. References**

[1] Allen, Colin & Wallach, Wendell & Smit, Iva. (2006). Why Machine Ethics?. Intelligent Systems, IEEE. 21. 12- 17. 10.1109/MIS.2006.83.

[2] Khari Johnson (2023). People Are Increasingly Worried AI Will Make Daily Life Worse, https://www.wired.com/story/fast-forward-people-are-increasingly-worried-artificial-intelligence/

[3] Johannes Himmelreic(2018). The everyday ethical challenges of self-driving cars. <https://theconversation.com/the-everyday-ethical-challenges-of-self-driving-cars-92710>.

[4] Jack Boeglin, *The Costs of Self-Driving Cars: Reconciling Freedom and Privacy with Tort Liability in Autonomous Vehicle Regulation*, 17 YALE J.L. & TECH. 171 (2015).

[5] Rayhan, Shahana. (2023). Ethical Implications of Creating AGI: Impact on Human Society, Privacy, and Power Dynamics. Artificial Intelligence Review.

[6] S. Khellat-Kihel, R. Abrishambaf, J.L. Monteiro, M. Benyettou, Multimodal fusion of the finger vein, fingerprint and the finger-knuckle-print using Kernel Fisher analysis,Applied Soft Computing,Volume 42,2016,Pages 439-447,ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2016.02.008>.

[7] Chien-Yao Wang and Alexey Bochkovskiy and Hong-Yuan Mark Liao (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. [arXiv:2207.02696](https://arxiv.org/abs/2207.02696)