Introducing FRASA: An End-to-End Reinforcement Learning Agent for Fall Recovery and Stand Up of Humanoid Robots

Marc Duclusaud[1]

Abstract— **Humanoid robotics faces significant challenges in achieving stable locomotion and recovering from falls in dynamic environments. Traditional methods, such as Model Predictive Control (MPC) and Key Frame Based (KFB) routines, either require extensive fine-tuning or lack real-time adaptability. This presentation introduces FRASA, a Deep Reinforcement Learning (DRL) agent that integrates fall recovery and stand up strategies into a unified framework. Leveraging the Cross-Q algorithm, FRASA significantly reduces training time and offers a versatile recovery strategy that adapts to unpredictable disturbances. Comparative tests on Sigmaban humanoid robots demonstrate FRASA superior performance against the KFB method deployed in the RoboCup 2023 by the Rhoban Team, world champion of the KidSize League.**

I. Introduction

Humanoid robotics has made significant strides in recent years, driven by advancements in both hardware and machine learning. One of the key challenges in this field is developing robots that can autonomously perform complex tasks, including locomotion, in dynamic and unpredictable environments. Reinforcement Learning (RL) has emerged as a powerful tool for addressing these challenges, enabling robots to learn from interactions with their environment and have more adaptive responses.

Achieving stable locomotion is one of the primary obstacles, if not the greatest, in humanoid robotics. The difficulty of this task arises primarily from the limited number of ground contact points for a humanoid, which results in an inherently unstable standing pose. Moreover, a high Center of Mass (CoM) increases the risk of damage in the event of a fall for humanoid robots. This leads researchers to prioritize fall prevention over recovery.

The proposed approach does not focus on the problem of push recovery but rather on **fall recovery**, which involves creating new points of contact with the ground as a strategy to regain stability in minimal time. The approach usually used in robotics competitions is to wait for the fall to end before starting a **stand up** routine, generally based on key frame animations. This method is simple to implement but lacks versatility, requires extensive fine-tuning, and must be frequently adjusted due to motor wear and loss of precision.

In this presentation, we introduce a Fall Recovery And Stand up Agent (**FRASA**) that integrates both fall recovery and stand up strategies into a single agent, aiming to resume walking as fast as possible after a disturbance. The adaptative response allowed by FRASA in the case of backwards and frontwards disturbances is presented in Fig. 1.

[1]Univ. Bordeaux, CNRS, LaBRI, UMR 5800, 33400 Talence, France.

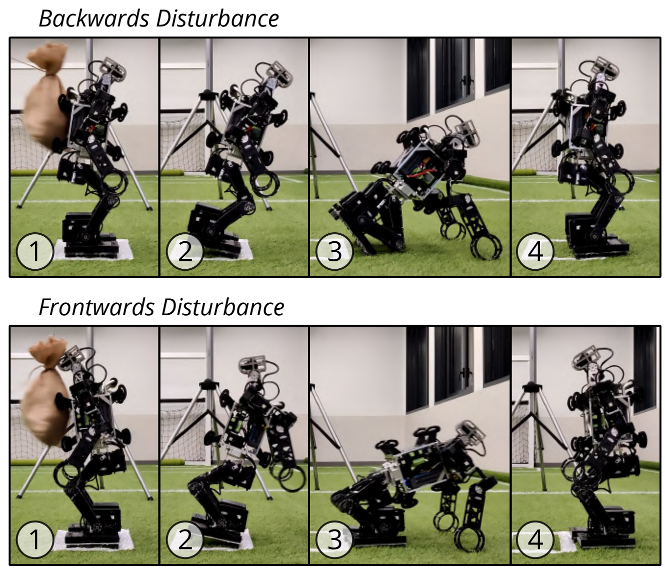

Fig. 1. FRASA adaptative response to a backwards and a frontwards disturbances on the Sigmaban platform. The recovery behavior using the arms accelerates the return to a stable position while minimizing the risk of damage.

The key contributions of FRASA are:

- **Unified task handling**: The proposed reward function enables the DRL agent to efficiently address both fall recovery and stand up tasks within a unified framework.
- **Reduced training time**: By leveraging the Cross-Q algorithm, FRASA achieves effective training in significantly less time compared to existing RL approaches.
- **Versatile and adaptative recovery strategy**: The use of a DRL agent allows adaptation to unpredictable real-world feedback and recovery from various postures without the need for expert tuning.

Tests are conducted on Sigmaban humanoid robots to compare FRASA with a Key Frame Based (KFB) approach. The KFB method used for comparison was deployed in the RoboCup 2023 by the Rhoban Team, world champion of the KidSize League. Real robots experiments are presented in the following video[1].

II. Experiments

To validate the performance of FRASA in a real environment, we deploys it on the Sigmaban robotic platform.

A. Stand Up Experiment

A first experiment is conducted to compare the ability to stand up from prone and supine positions using both a

[1]https://youtu.be/NL65XW0O0mk

KFB recovery process and FRASA. A KFB stand up method defines arbitrary key positions determined thanks to expert knowledge for the robot to pass through, interpolating motor positions between them to create a stand up trajectory. The KFB method used in the experiments was developed for RoboCup 2023 by the Rhoban Team and demonstrated both its effectiveness and efficiency.

The robot is placed on its back and face down, aiming to stand and begin walking. The metric studied is the time from movement initiation to the start of walking. To initiate walking, the error in the state vector components—comprising the 5 controlled joints and the trunk pitch—must remain below $5°$ for 0.5 seconds.

B. Fall Recovery Experiment

A second experiment compares the recovery ability of the KFB process and FRASA after being pushed from a static standing pose. The setup, shown in Fig. 2, uses a pendulum mechanism to release a mass from varying distances onto the robot in a repeatable manner. A cord allows retrieval of the weight post-impact, preventing interference during recovery.

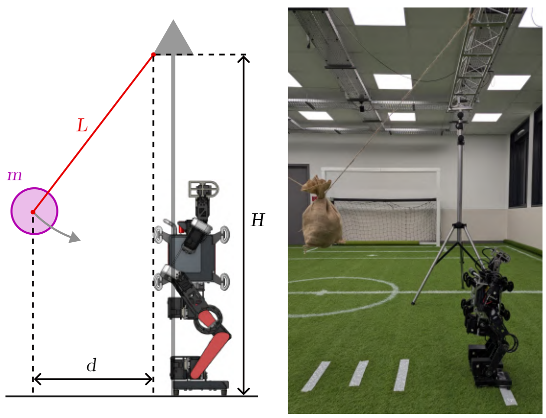

Fig. 2. Experimental setup for inducing repeatable disturbances

In this experiment, the robot is placed under the pendulum and subjected to repeated disturbances with varying impact intensities. After a disturbance, the robot waits to return to a stable position close to the target posture before beginning to walk in place.

The metric studied is the duration of disturbance rejection. Time measurement begins when any value in the state vector deviates by more than $5°$ from the neutral walking posture values. Once in this unstable state, all values in the state vector must have an error of less than $5°$ relative to the target posture for 0.5 s for the robot to be considered stable again and able to initiate walking. The time measurement ends at the beginning of the walking phase.

III. Results

The stand up experiment is conducted by performing the stand up task 20 times for each side and each method. The average stand up times are presented in Table I. We observe that FRASA outperforms the KFB method in both possible configurations: FRASA completes the stand up from the supine position in 53% and from the prone position in 68% of the time required by the KFB method. This corresponds to average gains of 2.38 seconds and 1.02 seconds, respectively.

TABLE I

Comparison of average stand up times for FRASA and KFB method from prone and supine positions

	Prone position		Supine position	
Method	FRASA	KFB	FRASA	KFB
Average Time / s	**2.135** ± 0.042	3.154 ± 0.005	**2.678** ± 0.178	5.06 ± 0.008

The standard deviation is represented using the notation $\pm$

The results of the fall recovery experiment are detailed in Table II. Each configuration of distance, method, and side is repeated 10 times, and the mean instability time is calculated. An estimate of the kinetic energy transferred at the moment of impact is calculated by considering the pendulum weight as a point mass. The weakest backwards disturbance does not cause any imbalance in the robot.

TABLE II

Comparison of average instability durations of FRASA and the KFB method after disturbances

	Frontwards Disturbance			Backwards Disturbance	
d / m	0.56	0.75	0.89	0.75	0.89
Energy / J	4.0	5.5	7.3	5.5	7.3
FRASA / s	0.54 ± 0.02	**2.41** ± 0.04	**2.44** ± 0.03	0.62 ± 0.10	**2.26** ± 0.11
KFB / s	0.57 ± 0.02	5.96 ± 0.11	5.74 ± 0.05	0.49 ± 0.10	4.47 ± 0.23

The standard deviation is represented using the notation $\pm$

In all configurations, FRASA achieves comparable or shorter instability times than the KFB method. The only configuration where the KFB method has a lower mean instability time than FRASA is for a 5.5J disturbance from the back. However, the standard deviation ranges for the two methods overlap, indicating that the difference is not statistically significant.

For the most significant impacts ($d = 0.89$m representing 7.3J), FRASA is able to reject the disturbance in 42% of the time required by the KFB method for a front impact and in 51% of the time for a back impact. FRASA also demonstrates superior performance for 5.5J frontwards disturbance, where the return to stability takes only 40% of the time compared to the KFB method.

Overall, FRASA demonstrates its superiority both in rejecting significant disturbances and in recovering from a prone or a supine position, surpassing the KFB method.

IV. Conclusion

Experiments on the Sigmaban platform demonstrate FRASA's superiority over the KFB method in both standing up and fall recovery, achieving these tasks with increased efficiency.

To further enhance FRASA's performance, a promising direction would be to finalize its training directly on the robot (online) to potentially reduce the sim-to-real gap. Additionally, improving the modeling of actuators in our simulator could help bridge this gap even further.