POLITECNICO DI MILANO

PhD in INGEGNERIA DELL'INFORMAZIONE / INFORMATION TECHNOLOGY - 40th cycle

Research Area n. 1 - Computer Science and Engineering

PNRR 630 Research Field: LARGE LANGUAGE MODELS TO ENHANCE AUTONOMOUS ROBOT BEHAVIORS

Monthly net income of PhDscholarship (max 36 months)
€ 1500.0 In case of a change of the welfare rates during the three-year period, the amount could be modified.

Context of the research activity	
Motivation and objectives of the research in this field	Robots have been successfully used in a wide variety of applications, ranging from heavy industry to precisemanufacturing, from agriculture to healthcare. However, many of these applications rely on structured environmentsor predefined tasks and directives. Such an approach is still necessary because robots lack a deeper understanding of the world and the ability to tackle unpredictable situations. Recently, large language models, such as GPT, Llama, and Gemini, have shown impressive capabilities in prompt understanding and text generation. In particular, they can successfully generate structured text (e.g., executable code,graphs) and manage common sense scenarios. For example, spatial relationships (i.e., a box in front of the sofa), qualitative comparisons (i.e., moving to the colder location), and composability (i.e., a stool has a seat and four legs). Recent works have shown how LLMs can be used to control robots in performing complex tasks. Such works include the generation of Python code, both directly and via the support of a template, the generation of intermediate descriptions to convert human prompts to executable actions, and the exploitation of LLMs to enhance the understanding of the environment. Additionally, some approaches went beyond the use of

	text by extending theinput to manage multiple data formats (e.g., text and images) or by adopting an end-to-end solution that outputs robot configurations. While these works achieve various degrees of success in multiple scenarios, they all have a significant downside whenused in robotics. LLMs, by definition, require a large amount of resources (i.e., memory and computational power) to operate. Specifically, there is a direct correlation between the size of the model and the quality of results achieved.This significant use of resources is incompatible with many robotic applications since robots have physical constraints that limit their computational power, such as battery consumption, size, and weight. This research proposal aims to overcome the limitations of large language models in robotics by exploring the capabilities of lightweight LLMs. Lightweight models are characterized by a limited number of parameters (e.g., less than twenty billion) making them deployable on consumer hardware or embedded computing devices. Therefore,they are compatible with the limited resources available to a robot. In particular, this work will focus on how multiplel ightweight LLMs can be combined to achieve a higher level of autonomy in the robot.
Methods and techniques that will be developed and used to carry out the research	In the context of this work, the main activity will be developing a system based on lightweight LLMs to generate and coordinate complex behaviors for robots. The initial phase of the research will be an in-depth analysis of the state-of-the-art about the interaction betweenLLMs and robotics. The study will focus on existing applications that exploit the capabilities of large models toenhance robotic behaviors. Particular attention will be given to works that rely on limited hardware while achievingsuccessful results. The analysis will extend beyond text-based models and include multi-modal solutions, such asvision language models and transformer-based architectures for robot control. The analysis of the state-of-the-art will lead to a selection of large models suitable to be used in robotic applications,such as general behaviors, planning, navigation, scene understanding, and human-robot interaction. From here, thefocus will shift from the single

	model to the interaction of multiple models. Independently, lightweight large models have limited capabilities compared to their counterparts with moreparameters. However, some work already demonstrated how the interaction of multiple models, especially withdifferent modalities, can lead to greater global performance. Therefore, in this research, significant work will be doneto coordinate and orchestrate multiple lightweight large models to achieve an autonomous robot capable of higherautonomy (i.e., without the need for continuous human prompting) and capable of adapting to highly dynamic environments.
Educational objectives	The PhD program offers the highest level of formation, leading to a high expertise in the chosen field and training inthe latest research topics currently explored by the scientific community. The candidate will acquire a wide set of skills regarding autonomous robots, and large language models.
Job opportunities	Doctoral graduates have opportunities both in academia and in industry. The collaboration with industrial partners allowed past PhD graduates to find satisfactory job positions in private companies. Opportunities in academia are also made possible.
Composition of the research group	2 Full Professors 0 Associated Professors 5 Assistant Professors 5 PhD Students
Name of the research directors	Matteo Matteucci

Contacts
matteo.matteucci@polimi.it, 02 2399 3470, https://www.deib.polimi.it/eng/people/details/267262

Additional support - Financial aid per PhD student per year (gross amount)	
Housing - Foreign Students	--
Housing - Out-of-town residents (more than 80Km out of Milano)	--

Scholarship Increase for a period abroad

Amount monthly	750.0 €
By number of months	6

National Operational Program for Research and Innovation	
Company where the candidate will attend the stage (name and brief description)	Oversonic Robotics S.r.l Società Benefit (Besana Brianza)
By number of months at the company	6
Institution or company where the candidate will spend the period abroad (name and brief description)	Pal Robotics (Barcelona, Spain)
By number of months abroad	6

Additional information: educational activity, teaching assistantship, computer availability, desk availability, any other information

EDUCATIONAL ACTIVITIES (purchase of study books and material, including computers, funding for participation in courses, summer schools, workshops and conferences): financial aid per PhD student.

TEACHING ASSISTANTSHIP: availability of funding in recognition of supporting teaching activities by the PhD student. There are various forms of financial aid for activities of support to the teaching practice. The PhD student is encouraged to take part in these activities, within the limits allowed by the regulations.

COMPUTER AVAILABILITY:
1st year: Yes
2nd year: Yes
3rd year: Yes