Welcome to Robophilosophy 2014

Since Giuseppe Veruggio coined the term "robo-ethics" in 2004, philosophers have observed with increasing attention the rise and progress of social robotics. Conferences in artificial intelligence, robotics, sociology, anthropology, and philosophy of technology included an increasing number of talks on the specific philosophical problems that arise with the creation of artificial agents with social intelligence and interactive capabilities.

However, the philosophical reflection of social robotics pertains to *all* systematic areas of philosophy, not only to ethics and philosophy of mind, but also to metaphysics, ontology, epistemology, philosophy of science, philosophical anthropology and aesthetics, philosophy of culture, political philosophy, and even to philosophy of religion. To our knowledge this conference is the first philosophical event where researchers from all major disciplines of philosophy come together to discuss sociable robots, and this is one historical particularity we wish to highlight by speaking of "robo-philosophy."

More importantly, the term "robo-philosophy" is to invite reflections about a possible turning point in the discipline of philosophy itself. Given that all aspects of human life constitutively involve social interactions, it is hardly surprising that social robotics is a topic for all areas of philosophy; but if social interactions are no longer distinctly human, philosophy may have to change its job description. Can or should philosophy continue to understand itself as the 'inquiry into the human condition' or will it become the 'general theory of interactions'?

Differently put, will philosophy become "robo-philosophy"? For it seems that social robotics puts *all* philosophical questions into a new key. So far philosophical questions have taken the general form 'what is X (truth, justice, morality, agency etc.)?' by inquiring about conditions C for something's being X—e.g., when is an action morally good, a decision rational, a society just, a belief true, an item beautiful? With the rise of social robotics, our concern in philosophy is not only how we can state conditions C in the simplest, most adequate, and illuminating way. Suddenly the question is also whether we can or should restrict C in such a way that they apply only to humans, or whether we should take the viewpoints (?) and capacities of artificial intelligent beings into account.

Robophilosophy puts the ever present question of human exceptionalism—foi for any philosophical topic—into a new key. This manifests itself in a number of core issues that run across the different systematic perspectives in the philosophy of social robotics. In the very foreground is the simulation issue: Do we have reason to assume that certain human capacities cannot be simulated, if so, is this a sense of simulation that we make sufficiently explicit as to make it philosophically relevant? A second systematically cross-cutting theme is the issue of conceptual augmentation: Should we try to subsume the new kind of quasi-social interactions among humans and machines under the conceptual distinctions that we have worked out, or should we develop new conceptual tools? The third theme, the issue of rational desirability, has greatest public attention: Should we, humans, want social robots;

will they benefit our goals as human beings? These three generic themes run across the 48 talks of this conference and in the course of our discussions we will likely discover several more.

The focus of our conference, the future of social relations, is but one topical area in the wide field of philosophical reflections on developments in social robotics. But it is likely to remain a central topic, even though a look at the projectable development in social robotics may suggest that our interactions with robotic agents with continuous processing of internet information soon will profoundly deviate from the model of human social interactions. We hope that our joint discussions on "robo-philosophy" will bring us one step further in exploring Sherry Turkle's suggestion that we are currently creating what later generations will identify (decry?) as the "robotic moment in human history."

Johanna Seibt, Raul Hakli, Marco Nørskov