Form 3

Abstract of Thesis

Name （BARAGLIA JIMMY CAMILLE）	
Title	Prediction Error Minimization for the Emergence of Prosocial Helping Behavior (予測誤差最小化に基づく向社会的援助行動の創発)

The aims of the work presented in this dissertation are to understand the emergence and the development of prosocial helping abilities in infants, especially helping behavior, and to reproduce them into robots. The main objectives of our studies are:

- To propose a general motivation mechanism that explains how infants may acquire prosocial helping abilities, and design a computational model to reproduce such behavior in robot.
- To construct a robotic system capable to fluently and efficiently interact with others in the context of joint tasks collaboration.

Our approach throughout our work is to create robotic models that build representations of their environment using statistical predictor based systems and perform actions that minimize prediction error. By doing so, our systems become capable to develop cognitive and prosocial abilities such as helping behavior or collaboration, similarly to what infants are capable to do. We performed two studies that describe major steps from the development of prosocial helping abilities to efficient human-robot collaboration.

- The first study presents a general mechanism explaining the emergence of prosocial helping behavior based on goal understanding. It is known that infants are capable to help others achieving goal that they are capable to understand. Traditional theory suggest that empathy and emotion contagion play capital roles in the motivation for early helping behavior. However, recent studies have shown that infants can similarly help simple animated shapes to achieve their goal, implying that another mechanism, more general, may be involve in the emergence of helping behavior. Authors claimed that infants may be primed to help by unfulfilled goal. We suggest that when observing actions, infants are capable to predict their goals based on self-experience and to perform actions if the goal is not achieved. In other words, that infants help others to minimize prediction error. To evaluate our hypothesis, we design a computational model based on psychological studies and implement it in real and simulated robots. Our experimental results demonstrate that our robots could spontaneously generate helping behavior by being motivated by the minimization of prediction error.
- The second study focuses on how to use previously described abilities to generate efficient and fluent human-robot interactions. We address the question of whether and when a robot should take initiative during collaborative human-robot task execution. We design and conceive a robotic system capable to autonomously perform table-top manipulation tasks while monitoring the environmental state and human activity. To evaluate our system, we implement three different initiative conditions to trigger the robot's actions. Human-initiated help gives control of the robot action timing to the user; robot-initiated reactive help triggers robot assistance when it detects that the human needs help; robot-initiated proactive help makes the robot help whenever it can. The result of our user study (N=18) give us significant proofs that proactive robots perform best, but are perceived less social than the human-initiated ones.

Together, these two studies describe mechanisms explaining how a robot can develop helping behavior based on prediction error minimization and perform efficient and fluent collaboration. In addition, our results contribute to the field of developmental science by proposing a general and likely mechanism for the development of prosocial helping abilities, and in particular for helping behavior.

様式 7

論文審査の結果の要旨及び担当者

氏　名	（ Baraglia Jimmy Camille ）		
論文審査担当者		（職）	氏　名
	主 査	教授	浅田 稔
	副 査	教授	石黒 浩
	副 査	教授	細田 耕
	副 査	特任准教授	長井 志江
	副 査	教授	中谷 彰宏
	副 査	教授	平田 勝弘
	副 査	教授	南埜 宜俊
	副 査	教授	菅沼 克昭

論文審査の結果の要旨

The aims of the work presented in this dissertation are to understand the emergence and the development of prosocial helping abilities in infants, especially helping behavior, and to reproduce them into robots. The main objectives of our studies are:

- To propose a general motivation mechanism that explains how infants may acquire prosocial helping abilities, and design a computational model to reproduce such behavior in robot.
- To construct a robotic system capable to fluently and efficiently interact with others in the context of joint tasks collaboration.

Our approach throughout our work is to create robotic models that build representations of their environment using statistical predictor based systems and perform actions that minimize prediction error. By doing so, our systems become capable to develop cognitive and prosocial abilities such as helping behavior or collaboration, similarly to what infants are capable to do. We performed two studies that describe major steps from the development of prosocial helping abilities to efficient human-robot collaboration.

- The first study presents a general mechanism explaining the emergence of prosocial helping behavior based on goal understanding. It is known that infants are capable to help others achieving goal that they are capable to understand. Traditional theory suggest that empathy and emotion contagion play capital roles in the motivation for early helping behavior. However, recent studies have shown that infants can similarly help simple animated shapes to achieve their goal, implying that another mechanism, more general, may be involve in the emergence of helping behavior. Authors claimed that infants may be primed to help by unfulfilled goal. We suggest that when observing actions, infants are capable to predict their goals based on self-experience and to perform actions if the goal is not achieved. In other words, that infants help others to minimize prediction error. To evaluate our hypothesis, we design a computational model based on psychological studies and implement it in real and simulated robots. Our experimental results demonstrate that our robots could spontaneously generate helping behavior by being motivated by the minimization of prediction error.

- The second study focuses on how to use previously described abilities to generate efficient and fluent human-robot interactions. We address the question of whether and when a robot should take initiative during collaborative human-robot task execution. We design and conceive a robotic system capable to autonomously perform table-top manipulation tasks while monitoring the environmental state and human activity. To evaluate our system, we implement three different initiative conditions to trigger the robot's actions. Human-initiated help gives control of the robot action timing to the user; robot-initiated reactive help triggers robot assistance when it detects that the human needs help; robot-initiated proactive help makes the robot help whenever it can. The result of our user study (N=18) give us significant proofs that proactive robots perform best, but are perceived less social than the human-initiated ones.

Together, these two studies describe mechanisms explaining how a robot can develop helping behavior based on prediction error minimization and perform efficient and fluent collaboration. In addition, our results contribute to the field of developmental science by proposing a general and likely mechanism for the development of prosocial helping abilities, and in particular for helping behavior.

以上のように、本論文は乳幼児における向社会的援助行動の創発について，計算論的視点から予測誤差最小化に基づく新たなメカニズムを提案し，さらに，人とロボットの援助行動を介したインタラクションの設計においても，予測誤差最小化のメカニズムが有効であることを被験者実験により証明している．よって本論文は博士論文として価値あるものと認める．