

ABSTRACT

“The question is not whether the intelligent machines can have any emotions, but whether machines can be intelligent without any emotions”.

Minsky in *The society of mind* (1986).

The objective of this thesis is to develop a decision making system for an autonomous and social robot. This system is composed by several subsystems: a motivational system, a drives system and an evaluation and behaviour selection system. All systems are based on motivations, drives and emotions. These concepts are described in detail in this thesis.

Due to the difficulties of working with a real robot, it has been decided to implement this decision making system on virtual agents as a previous step. These agents live in a virtual world which has been built using a text based MUD (Multi-User Domain). In this world the agents can interact with each other, allowing the social interaction, and with the other objects present in the world. The reason why this text based game was selected, instead of a modern one with graphic interfaces, is that the interpretation of the information is much simpler.

The selection of behaviours is learned by the agent using reinforcement learning algorithms. When the agent is not interacting with other agents he uses the Q-learning algorithm. When social interaction exists, the rewards the agent receives depend not only on his own actions, but also on the action of the other agent. In this case, the agent uses multi-agent learning algorithms, also based on Q-learning.

The inner state of the agent is considered as the dominant motivation. This fact causes that the system is not completely markovian and therefore, not so easy to work with. In order to simplify the learning process, the states related to the objects are considered as independent among them. The state of the agent is a combination between his inner state and his state in relation with the rest of agents and objects.

The fact that the states in relation to the objects are considered as independent, causes that the possible relation between objects is ignored. In fact, the actions related to an object, may affect the state in relation to the other objects, causing “collateral effects”. In this thesis, a new variation of Q-learning is proposed to consider these effects.

This system uses happiness and sadness as positive and negative reinforcement respectively. Therefore, behaviours are not going to be selected to satisfy the goals determined by the motivations of the agent, but to reach happiness and avoid sadness.

The appraisal emotions theories state that emotions are the result of evaluation processes and therefore they are subjective. Based on those theories, it is going to be considered, in this decision making system, that certain emotions are going to be generated from the evaluation of the wellbeing of the agent. The wellbeing measures how much the needs of the agent are satisfied. Happiness is produced because something good has happened, i.e. an increment of the wellbeing is produced. On the contrary, sadness is produced because something bad has happened, so the wellbeing decreases.

Finally, another emotion is introduced: Fear. Fear is presented from two points of view: to be afraid of executing risky actions, or to be afraid of being in a dangerous state. In this last case, fear is considered as a motivation, in accordance with other emotions theories.

RESUMEN

“La cuestión no es si las máquinas inteligentes pueden tener alguna emoción, sino si las máquinas pueden ser inteligentes sin ninguna emoción”.

Minsky en *La sociedad de la mente* (1986).

El objetivo de esta tesis es desarrollar un sistema de toma de decisiones para un robot personal social y autónomo. Este sistema consiste en varios módulos: el módulo motivacional, el módulo de los drives y el módulo de valoración y selección de comportamientos. Estos módulos están basados en motivaciones, drives y emociones, conceptos que son ampliamente desarrollados en esta tesis.

Debido a las dificultades de trabajar con un robot real, como primer paso en esta investigación, se ha optado por la implementación previa de esta arquitectura en agentes virtuales. Dichos agentes viven en un mundo virtual que ha sido construido utilizando un juego de rol basado en texto del tipo MUD (Multi User Domain). En este mundo los agentes interactúan entre sí y con diferentes objetos. Se ha elegido este tipo de juegos basados en texto en lugar de otros más modernos con interfaces gráficas, porque la interpretación de la información es mucho más sencilla.

La selección de los comportamientos es aprendida por el agente mediante la utilización de algoritmos de aprendizaje por refuerzo. Cuando el agente no interactúa con otros agentes utiliza el algoritmo Q-learning. Cuando existe interacción social el refuerzo que recibe el agente es debido a la acción conjunta con el otro agente y, en este caso, hace uso de algoritmos de aprendizaje multi-agente, también basados en Q-learning.

El estado interno del agente se considera como la motivación dominante, lo que provoca que el sistema no sea completamente markoviano y por lo tanto no tan sencillo de tratar. Con el fin de simplificar el proceso de aprendizaje, los estados relacionados con los objetos del mundo se consideran como independientes entre sí. El estado del agente es una combinación de su estado interno y su estado en relación con el resto de agentes y objetos.

El hecho de considerar los estados en relación a los objetos como independientes, hace que no se tenga en cuenta la posible relación entre ellos. En realidad, las acciones

que se realizan con un objeto pueden llegar a afectar al estado en relación con los otros objetos, provocando “efectos colaterales”. En esta tesis se propone una variación original del algoritmo Q-learning que considera dichos efectos.

En este sistema se utilizan la felicidad y la tristeza como refuerzo positivo y negativo respectivamente. Por lo tanto, los comportamientos no son seleccionados para satisfacer los objetivos determinados por las motivaciones del agente, sino para alcanzar la felicidad y evitar la tristeza.

Las teorías de valoración de las emociones afirman que éstas son el resultado de procesos evaluadores y por lo tanto subjetivos. En base a estas teorías se considera, en este sistema de toma de decisiones, que ciertas emociones son generadas a partir de la valoración del bienestar del agente. Este bienestar mide el grado de satisfacción de las necesidades del agente. De esta forma se asume que la felicidad se produce cuando le sucede algo bueno al agente, aumentando el bienestar. La tristeza, por el contrario, se produce cuando pasa algo malo, provocando una disminución del bienestar.

Finalmente, se introduce la emoción del miedo desde dos puntos de vista: tener miedo a realizar acciones arriesgadas y tener miedo a estar en estados peligrosos. Para ello se considera el miedo como otra motivación del agente, lo cual también coincide con otras teorías de emociones.