

# Resumen

La *argumentación* es un modelo de razonamiento basado en la construcción y evaluación de argumentos que interactúan entre sí. Estos argumentos tienen como fin, soportar, explicar o atacar enunciados que pueden ser opiniones, decisiones, etc.

La argumentación ha sido utilizada en diferentes dominios, como razonamiento no-monótono, manejo de inconsistencias en bases de conocimiento y modelado de diferentes tipos de diálogos, en particular, persuasión y negociación. Un enfoque de negociación basado en argumentación tiene la ventaja de que además de intercambiar ofertas, se intercambian razones que las soportan y que eventualmente podrían llevar al receptor a cambiar sus preferencias. En consecuencia, se puede llegar a un acuerdo entre las partes de manera más conveniente.

De esta manera, adoptar un enfoque de este estilo en los problemas de decisión, tendría el beneficio de que el tomador de decisiones además de tener una buena elección, tendría las razones subyacentes que la soportan de una manera fácil de entender. La toma de decisiones basada en argumentación, es más afín a la forma en que los seres humanos deliberan y finalmente toman o entienden una elección.

La idea de articular decisiones en base a argumentos es relevante para varios enfoques diferentes de decisión, tales como decisión bajo incertidumbre, decisión multi-criterio y decisiones basadas en reglas. En general, puede decirse que los diversos enfoques propuestos construyen para cada decisión posible un conjunto de argumentos a su favor y un conjunto de argumentos en su contra. No obstante, los tomadores de decisiones no se limitan solamente a construir estos conjuntos, sino que ellos intercambian argumentos, los hacen interactuar, atacándose o defendiéndose entre ellos. Asimismo, dentro de un enfoque particular de decisión, existen propuestas muy diversas que difieren en: los lenguajes de representación de conocimiento y razonamiento que utilizan, si su definición está dada

dentro de una arquitectura de agente particular, la formalidad con la que se relacionan con enfoques clásicos de toma de decisiones, etc.

En particular, en esta Tesis, se presenta un framework original de toma de decisiones individuales que está vinculado formalmente con la Teoría de Decisión Clásica, pero justificado desde el punto de vista de la argumentación.

El framework propuesto para la toma de decisiones basadas en argumentación, combina el uso de argumentos y reglas de decisión. El criterio de comparación de argumentos utilizado, se basa en prioridades entre literales que representan las preferencias del agente (tomador de decisiones). Estos literales forman parte del programa del agente, y al realizar cambios en las relaciones de preferencia entre los mismos, se puede cambiar de manera flexible las preferencias del agente y el criterio de preferencia entre argumentos.

Nuestro enfoque también incluye una metodología simple para desarrollar las componentes de decisión del agente. De esta manera, un framework de decisión desarrollado siguiendo esta metodología exhibe propiedades interesantes, con respecto a la coherencia que muestra el agente en la toma de decisiones. Si el agente, tiene disponible todo el conocimiento relevante acerca de sus preferencias con respecto a todas las posibles alternativas que se le podrían presentar, entonces el comportamiento de elección correspondiente al framework propuesto, coincide con el comportamiento óptimo respecto de una *relación de preferencia racional*.

En cambio, si el agente tiene conocimiento parcial acerca de sus preferencias, las decisiones que éste tomará también exhibirán un cierto grado de coherencia. En particular en este caso, nuestro enfoque implementa una estructura de elección que satisface el *axioma débil de la preferencia revelada*.

Los principios enunciados en esta Tesis, serán ejemplificados en dos dominios de aplicación distintos. En el primero, se analiza cómo se desarrollan los procesos de decisión de un robot que realiza tareas de limpieza y debe decidir qué caja transportará a continuación. Se presentan ejemplos donde el robot posee información completa e incompleta acerca de sus preferencias de elección. El otro dominio, consiste de un agente de software de impresión que debe decidir a qué impresora de la red manda un cierto trabajo, en función de las características de las impresoras y de su disponibilidad. En este dominio particular, también se presentan ejemplos con información completa sobre las preferencias del agente y se muestra cómo desarrollar los componentes de decisión para la aplicación del framework.

# Abstract

*Argumentation* is a reasoning model based on the construction and evaluation of interacting arguments. These arguments are intended to support, explain or attack statements that can be decisions, opinions, etc.

Argumentation has been used in different domains, such as non-monotonic reasoning, handling inconsistency in knowledge bases and modeling different kinds of dialogues, in particular, persuasion and negotiation. An argumentation-based approach to negotiation has the advantage of exchanging in addition to offers, reasons that support these offers. Eventually, these reasons might lead their receivers to change their preferences. Consequently, an agreement may be more easily reached with such approach.

In this way, adopting such an approach in a decision problem would have the benefits of not only providing to the decision maker a good choice, but also the reasons underlying this recommendation, in a format that is easy to grasp. Argumentation-based decision making is more akin with the way humans deliberate and finally make or understand a choice.

The idea of articulating decisions on the basis of arguments is relevant to different decision problems or approaches such as decision making under uncertainty, multiple criteria decisions and rule-based decisions. Generally speaking, it can be stated that the different existing approaches build arguments in favor or against for each candidate decision. In fact, decision makers are not only limited to build these sets of arguments, they exchange arguments, make them interact, because attacks and rebuttals occur among them. Furthermore, within a particular decision approach, different proposals exist that mainly differ in: the reasoning and knowledge representation languages they use, the formality of the way they relate to classical approaches to decision making, if their definitions are formulated with respect to a particular agent architecture, etc.

In particular, in this Thesis, a novel framework to Individual Decision Making is presented. This framework is formally related to Classical Decision Theory but it is justified from an argumentative point of view.

The proposed framework for argumentation-based decision making combines decision rules and arguments. The comparison criterion used, is based on priorities among literals that represent the agent's preferences. These literals belong to the agent's program and modifying the preference relations among them, allows us to change the agent's decision policy in a flexible way.

Our approach includes a simple methodology for developing the decision components of the agent. In this way, a decision framework designed with this methodology, exhibits some interesting properties with respect to the coherence showed by the agent in making decision. If the agent (decision maker) has available all the relevant knowledge about its preferences among the different alternatives that could be conceivably posed to it, then its choice behavior coincides with the optimal one according to *rational preference relation*.

In opposition, if the agent has partial knowledge about its preferences, the decisions made by the agent still exhibit a behavior consistent with the *weak axiom of revealed preference* of the *choice-based approach*, a more flexible approach to Individual Decision Making than the *preference-based approach*.

The principles stated in this Thesis are exemplified in two different application domains. The first one consist of a robot that should make decisions about which box must be transported next while it performs cleaning tasks. Examples with complete and incomplete information about its preferences are presented, and it is mainly analyzed how the decision processes are developed. In the second domain, a printer software agent should decide to which printer on the LAN must send a certain job, depending on the printers' features and their availability. In this particular domain, examples with complete information about the agent's preferences are also presented, emphasizing how to develop the decision components of the framework to model this domain.