Decision Support Systems

Introduction

Andrea Campagner, PhD, Post-Doc Researcher

IRCCS Istituto Ortopedico Galeazzi, Milan, Italy
MUDI Lab, Department of Computer Science, University of Milano-Bicocca, Milan, Italy

Contacts: onyris93@gmail.com; andrea.campagner@unimib.it
https://andreacampagner.github.io/

5 April 2023

Table of Contents

Logistics

Important!

- Lessons begin at 8.45 (15 slack on the official 8.30);
- 15 minutes break each 45 minutes of lesson (e.g., if lesson is 8.30-10.30, then we start at 8.45, break at 9.30, restart at 9.45, then continue till 10.30);
- If you have any questions, feel free to interrupt!
- Content **will be** in the exam!

Requirements

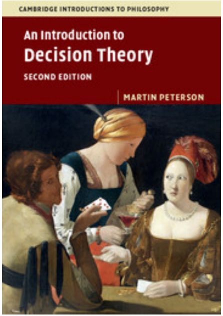
Requirements

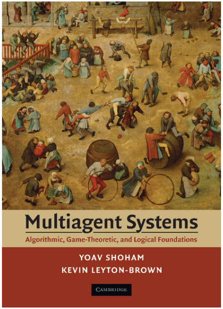
- Basics of computer science (algorithms, complexity theory)
- Basics of AI and ML
- Mathematical maturity (probability theory, analysis, linear algebra)

Don't be scared, course will be (mostly) self-contained!

Materials

Course will (mostly) follow the following books:





First is more basic (I will stick mostly to it), second is more advanced/mathematical (if you're interested in theoretical/computational issues)... we will also refer to some papers (made available on Moodle).

Introduction to the Course

Objective

Giving you a broad introduction to **(Normative) Decision Theory** and its applications in AI/ML

Introduction to the Course (cont.)

- But what is decision theory?
- Mathematical study of how **agents** *make* or *should make* **decisions** in a **decision-making setting**.

Introduction to the Course (cont.)

- Mathematical study of how **agents** *make* or *should make* **decisions** in a **decision-making setting**.
  - **Agent**: *Someone or something* that can *act upon something* else to obtain an *outcome*;

Introduction to the Course (cont.)

- Mathematical study of how **agents** *make* or *should make* **decisions** in a **decision-making setting**.
  - **Agent**;
  - **Decision**: *choice* of *one* among some *alternatives*;

Introduction to the Course (cont.)

- Mathematical study of how **agents** *make* or *should make* **decisions** in a **decision-making setting**.
  - **Agent**;
  - **Decision**;
  - **Decision-making setting**: Agent(s) + Environment + Decisions + Possible outcomes (+ possibly other things: e.g., information representation)

Introduction to the Course (cont.)

Distinction between *make* and *should make* is crucial

Introduction to the Course (cont.)

Distinction between *make* and *should make* is crucial

- Make: *Descriptive Decision Theory* $\rightarrow$ We are interested in studying and modeling how real agents make decisions in the real world;

Introduction to the Course (cont.)

Distinction between *make* and *should make* is crucial

- Make: *Descriptive Decision Theory*;
- Should Make: *Normative Decision Theory* $\rightarrow$ Assumption-driven endeavor, we are interested in studying what different assumptions (about the decision-making setting) entail about the *optimal* behavior of agents in that setting

Introduction to the Course (cont.)

Distinction between *make* and *should make* is crucial

- Make: *Descriptive Decision Theory*;
- Should Make: *Normative Decision Theory*.

In this part of the course we will focus on **Normative Decision Theory** (henceforth, only Decision Theory)... some Descriptive Decision Theory in the part of Prof. Cabitza!

Introduction to the Course (cont.)

Decision theory is a very broad field... different sub-fields depending on the structure of the decision-making setting!

Introduction to the Course (cont.)

Decision theory is a very broad field... different sub-fields depending on the structure of the decision-making setting!

- Single agent vs environment: **Decision Theory**;
- Multiple (self-interested) agents: **(Non-cooperative) Game Theory**;
- Multiple agents working together: **Coalitional Game Theory**;
- Many agents and one central planner: **Social choice theory**...

... and many others (bandit theory, reinforcement learning, ...)!

Introduction to the Course (cont.)

- Lots of applications: economics, computer science, artificial intelligence, machine learning, biology...
- We will view some recent/relevant applications in ML!
  - Decision Theory: ML models evaluation (net benefit theory);
  - Non-cooperative Game Theory: generative ML and GANs;
  - Coalitional Game Theory: explainable AI and SHAP;
  - Social Choice Theory: ensemble learning

Introduction to Decision Theory: Basics

Decision Theory: single agent vs environment... abstract approach

- Agent: set of *actions* $A = \{a_1, \ldots, a_n\}$
- Environment: set of *states* $S = \{s_1, \ldots, s_m\}$

The agent taking an action a_i when the environment is in a given state s_j determines an outcome o

- Outcome function: $O : A \times S \to \{o_{1,1}, o_{1,2} \ldots, o_{n,m}\}$

Introduction to Decision Theory: Decision Matrix

We can describe a decision-making setting through a **decision matrix**

Actions \ States	s_1	$\dots$	s_m
a_1	$o_{1,1}$	$\dots$	$o_{1,m}$
$\vdots$		$\ddots$	
a_n	$o_{n,1}$	$\dots$	$o_{n,m}$

Introduction to Decision Theory: Decision Matrix

We can describe a decision-making setting through a **decision matrix**... an example

	Fire	No fire
Insurance	No house and 100000€	House
No insurance	No house and 100€	House and 100€

$A = \{\text{Insurance}, \text{No insurance}\}$, $S = \{\text{Fire}, \text{No Fire}\}$

Introduction to Decision Theory (cont.)

Two main versions of Decision Theory:

- *No information* about the likelihood of the states: **Decision under Ignorance**
- Likelihood of the states is quantified by a *probability distribution*: **Decision under Risk**

As we will see, this makes a big difference!

Introduction to Decision Theory: Rationality

As we said, Decision Theory is an assumption-driven endeavor... our main assumption will be **rationality**

- There exists a pre-order P over the set of outcomes $O(A, S)$
  - Reflexivity: $\forall o, o \leq_P o$;
  - Transitivity: $\forall o_i, o_j, o_k \; o_i \leq_P o_j \wedge o_j \leq_P o_k \implies o_i \leq o_k$

Introduction to Decision Theory: Rationality

As we said, Decision Theory is an assumption-driven endeavor... our main assumption will be **rationality**

- There exists a pre-order P over the set of outcomes $O(A, S)$

P represents the preferences of the agent among outcomes:

- $o_i \leq_P o_j$ means that o_j is **(weakly) preferred to** o_i;
- $o_i <_P o_j$ (i.e., $o_i \leq_P o_j \wedge o_j \not\leq_P o_i$) means that o_j is **(strongly) preferred to** o_i;

Introduction to Decision Theory: Rationality (cont.)

As we said, Decision Theory is an assumption-driven endeavor... our main assumption will be **rationality**

- There exists a **preference** pre-order P over the set of outcomes $O(A, S)$

We will also typically assume that P is a *linear order*

- Completeness: $\forall o_i, o_j$ either $o_i \leq_P o_j$ or $o_j \leq_P o_i$;

This implies that we can define a function $U_P : O(A, S) \rightarrow \mathbb{R}$ s.t. $U_P(o_i) \leq U_P(o_j)$ iff $o_i \leq_P o_j$

U_P is usually called a *utility function* and it is the central notion in Decision Theory

Introduction to Decision Theory: Rationality (cont.)

As we said, Decision Theory is an assumption-driven endeavor... our main assumption will be **rationality**

- There exists a **preference** pre-order P over the set of outcomes $O(A, S)$
- There exists a **utility function** $U_P : O(A, S) \to \mathbb{R}$ s.t. $U_P(o_i) \leq U_P(o_j)$ iff $o_i \leq_P o_j$

The agent is rational if he/she acts so as to **maximize its utility**...
Consequence: we only actually care about utilities, not really about outcomes!

Introduction to Decision Theory: Rationality (cont.)

Two examples:

	Fire	No fire
Insurance	1	4
No insurance	-100	5

	Sixth egg is rotten	Sixth egg is fine
Add sixth egg	0	1
Do not add	0.83	0.83

Introduction to Decision Theory: Rationality (cont.)

Important Remark!

We only said that there exists a utility function U_P that encodes the preferences of the agent... But we have said nothing about its *scale*:

- Ordinal: values are arbitrary, only ordering matters;
- Cardinal: values (and their distances) matter

The scale is important as it defines which transformations on utilities are admissible... but we won't focus too much on this!

Decision under Ignorance

- As we mentioned previously, in **Decision under Ignorance** the agent has no information on the likelihood of the states...
- ...only knows which states could occur and which outcomes they determine!
- How a **rational** agent should make decisions in this setting?

Decision under Ignorance: Dominance

Central notion: **Dominance**

- $a_i \leq a_j$ (**weakly dominates**) if $\forall s \in S, U(O(a_i, s)) \leq U(O(a_j, s))$;
- $a_i < a_j$ (**strongly dominates**) if $\forall s \in S, U(O(a_i, s)) \leq U(O(a_j, s))$ and $\exists s_l \in S, U(O(a_i, s)) < U(O(a_j, s))$;

A rational agent **should never** consider a dominated action!

Decision under Ignorance: Dominance (cont.)

	Good chef	Bad chef
Monkfish	4	1
Hamburger	3	3
No main course	2	2

Choosing *No main course* is dominated by choosing *hamburger*: no matter the actual state of the environment (even if we do not know a priori), choosing *hamburger* is always better!

Decision under Ignorance: Dominance (cont.)

	Good chef	Bad chef
Monkfish	4	1
Hamburger	3	3
No main course	2	2

Dominance is a **very weak** decision rule: in most cases, it does not allow the agent to make a decision... e.g. *Monkfish* and *Hamburger* are not dominated, but clearly the agent considers them differently!

Decision under Ignorance: Decision Rules

- This means that we have to make some other assumptions other than rationality;
- Each (set of) assumption determines a **decision rule**: infinitely many rules, we only care that they are coherent with dominance!

Decision under Ignorance: Maximin

The agent wants to maximize utility in the worst case

	Bacterial infection	Viral infection	Stress
Take antibiotics	1	*-1*	*-1*
Take anti-fever	0.5	0.5	*-0.5*
No medication	*-1*	-1	0

It is the behavior of a very risk-averse agent: we focus on the worst possible case

Decision under Ignorance: Maximax

The agent wants to maximize utility in the best case

	Bacterial infection	Viral infection	Stress
Take antibiotics	*1*	-1	-1
Take anti-fever	*0.5*	*0.5*	-0.5
No medication	-1	-1	*0*

It is the behavior of a very optimistic agent: we focus on the best possible case

Decision under Ignorance: Averaging

- Clearly, maximin and maximax are very extreme: either maximally optimistic or maximally pessimistic...
- Obviously, we can do something in-between: we assign weights to the outcomes and average them (if interested, these are called **Order-weighted Average** operators);

Weights do not reflect in any way probabilities, and you should not think of them as such... in Decision under Ignorance we have no probabilities!!!

Decision under Ignorance: Minimax Regret

A variation on maximin, in which we consider the **regret**

- $Regret(a, s) = U(O(a, s)) - \max_{a' \in A} U(O(a', s))$
- The agent makes decisions in order to minimize the regret

	Bacterial infection	Viral infection	Stress
Take antibiotics	0	*-1.5*	-1
Take anti-fever	*-0.5*	*0*	*-0.5*
No medication	*-2*	-1.5	0

Regret is a **loss function**: we want to minimize the loss function... very popular in ML!

Decision under Ignorance: Indifference Principle

A way to transform a Decision under Ignorance problem into a Decision under Risk one

- We assign to each state s a probability of $\frac{1}{|S|}$
- We then make decisions according to the **expected utility maximization** rule (more on this later)

	Bacterial infection	Viral infection	Stress	EU
Take antibiotics	1	-1	-1	$-\frac{1}{3}$
Take anti-fever	0.5	0.5	-0.5	0.167
No medication	-1	-1	0	- 0.667

This approach is very popular but contested: if we have no information about probabilities, why are we using them?

Decision under Risk

- In the case of Decision under Ignorance we have seen many different rules... no single best one!
- But some rules are more popular: maximin (game playing), OWA operators (operations research), minimax regret (game playing, ML)
- The situation is very different in the case of Decision under Risk... One single rule: **Expected Utility Maximization**!

Decision under Risk: Probability

We assume that the agent quantifies the likelihood of the states using a **probability distribution**

- $p : S \to [0, 1]$, s.t. $\sum_{s \in S} p(s) = 1$

We won't focus on the semantics of the distribution, but two main ones:

- **Frequentist**: p encodes the actual likelihood of the states;
- **Subjective**: p encodes the belief of the agent about the likelihood of the states

Decision under Risk: Expected Utility

We assume that the agent quantifies the likelihood of the states using a **probability distribution**

- $p : S \to [0, 1]$, s.t. $\sum_{s \in S} p(s) = 1$

The distribution p defines the **expected utility** of an action a

- $EU(a) = \sum_{s \in S} p(s)U(O(a, s))$

Intuitively, represents how many units of utility the agent expects to gain on average

Important Remark

EU requires that the utility scale is cardinal: averaging makes no sense when the scale is ordinal!!!

Decision under Risk: Expected Utility

Expected Utility Maximization requires that the agent selects the action that maximizes the Expected Utility

	Bacterial infection (0.05)	Viral infection (0.15)	Stress (0.8)	EU
Take antibiotics	1	-1	-1	-0.9
Take anti-fever	0.5	0.5	-0.5	-0.3
No medication	-1	-1	0	-0.2

Von Neumann-Morgenstern Theorem

A rational agent should make decisions according to Expected Utility Maximization