هوش مصنوعی

فصل ۱۶

اتخاذ تصمیم‌های ساده

Making Simple Decisions

کاظم فولادی قلعه
دانشکده مهندسی، دانشکدگان فارابی
دانشگاه تهران

http://courses.fouladi.ir/ai

عامل‌های نظریه‌تصمیمی

DECISION-THEORETIC AGENTS

عامل نظریه‌تصمیمی
Decision-Theoretic Agents

عاملی که تصمیم‌های رسیونال را
بر اساس باورها و خواسته‌هایش اتخاذ می‌کند.

متغیرهای تصمیم‌گیری

باورهای عامل	خواسته‌های عامل
What agent believes	*What agent wants*

شرایط تصمیم‌گیری

عدم اطمینان	هدف‌های متناقض
Uncertainty	*Conflicting goals*

انواع تصمیم

انواع تصمیم	
تصمیم ساده *Simple Decision*	**تصمیم پیچیده** *Complex Decision*
تصمیم‌های تک‌ضرب *One-Shot*	**تصمیم‌های چندمرحله‌ای** *Multi-Stage*
تصمیم بر روی یک کنش	تصمیم بر روی یک دنباله از کنش‌ها

نظریه‌ی تصمیم

نظریه‌ی عمومی تصمیم‌های رسیونال

DECISION THEORY

نظریه‌ی تصمیم *Decision Theory*	=	نظریه‌ی سودمندی *Utility Theory*	+	نظریه‌ی احتمال *Probability Theory*

یک عامل رسیونال است

اگر و فقط اگر

کنشی با بالاترین سودمندی مورد انتظار را انتخاب کند.

(متوسط آماری) میانگین روی همه‌ی برآمدهای ممکن آن کنش

اصل **حداکثر امید سودمندی**
maximum expected utility (MEU)

۱

ترکیب باورها و مطلوبیت‌ها تحت عدم‌اطمینان

انواع تصمیم

تصمیم‌های ساده

انواع تصمیم	
تصمیم ساده *Simple Decision*	تصمیم پیچیده *Complex Decision*
تصمیم‌های تک‌ضرب *One-Shot*	تصمیم‌های چندمرحله‌ای *Multi-Stage*
تصمیم بر روی یک کنش	تصمیم بر روی یک دنباله از کنش‌ها

سودمندی

تابع سودمندی

UTILITY

ترجیح‌های عامل بین حالت‌های دنیا، توسط مفهوم سودمندی مشخص می‌شود.
preferences

تابع سودمندی

حالت (دنباله‌ی حالت‌های محیط) را به یک عدد حقیقی نگاشت می‌دهد.

$$U : S^* \to \mathbb{R}$$

$$s \mapsto U(s)$$

$U(s)$ عدد سودمندی، میزان مطلوبیت یک حالت را بیان می‌کند.
desirability

امید سودمندی

سودمندی مورد انتظار

EXPECTED UTILITY

A یک کنش غیرقطعی که حالت‌های برآمد آن عبارت است از: $Result_i(A)$

E شاهد موجود عامل درباره‌ی دنیا

$Do(A)$ گزاره‌ی: «کنش A در حالت فعلی اجرا می‌شود.»

$EU(A|E)$ امید سودمندی کنش به شرط شاهد داده‌شده

سودمندی مورد انتظار

$$EU(A|E) = \sum_i P(Result_i(A)|Do(A), E)U(Result_i(A))$$

ماکزیمم امید سودمندی

اصل **حداکثر امید سودمندی**

maximum expected utility (MEU)

یک عامل رسیونال باید کنشی را انتخاب کند که
امید سودمندی آن را ماکزیمم نماید.

$$action = \underset{a}{\operatorname{argmax}} \, EU(a \,|\, e)$$

اصل *MEU*: کنشی را انتخاب کنید که امید سودمندی را ماکزیمم می‌کند

نکته: یک عامل می‌تواند کاملاً رسیونال باشد (سازگار با MEU)،
حتی بدون یک بازنمایی یا کار با سودمندی‌ها و احتمالات
(مثلاً با استفاده از جدول مراجعه برای بازی TicTacToe)

اتخاذ تصمیم‌های ساده

۲

پایه‌های نظریه‌ی سودمندی

نظریه‌ی سودمندی

UTILITY THEORY

نظریه‌ی سودمندی *Utility Theory*	نظریه‌ای برای بازنمایی و استدلال برای ترجیح‌ها

این نظریه بیان می‌کند که هر حالت برای یک عامل درجه‌ای از سودمندی را دارد
و عامل حالتی که سودمندی بیشتری دارد را ترجیح می‌دهد.

عامل باید بین **برآمدهای** ممکن مختلف طرح‌های گوناگون، **ترجیح‌هایی** داشته باشد.
outcome **preferences**

نظریه‌ی سودمندی

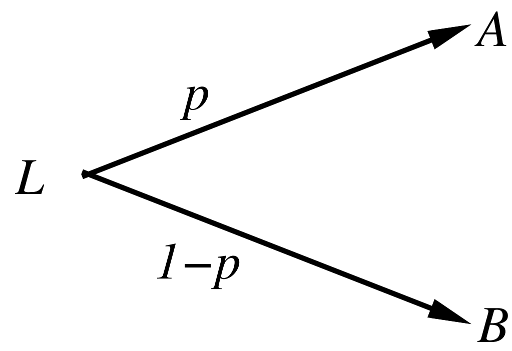
ترجیح‌ها

PREFERENCES

یک عامل می‌خواهد بین جایزه‌ها $(A,\ B,\ \ldots)$ در یک لاتاری انتخاب کند.

لاتاری = وضعیتی با جایزه‌های نامطمئن

Lottery $L = [p, A;\ (1-p), B]$



نمادگذاری

A preferred to B	$A \succ B$	A به B ترجیح داده می‌شود.
indifference between A and B	$A \sim B$	A با B تفاوتی ندارد.
B not preferred to A	$A \succsim B$	B به A ترجیح داده نمی‌شود.

نظریه‌ی سودمندی

ترجیح‌های رسیونال

Rational Preferences

ترجیح‌ها در یک عامل رسیونال باید از قیدهایی تبعیت کنند.

ترجیح‌های رسیونال ⇐ رفتار قابل توصیف به‌عنوان ماکزیمم‌سازی امید سودمندی

قیدهای ترجیح‌های رسیونال	
Orderability	ترتیب‌پذیری
$(A \succ B) \vee (B \succ A) \vee (A \sim B)$	
Transitivity	تراگذری
$(A \succ B) \wedge (B \succ C) \;\Rightarrow\; (A \succ C)$	
Continuity	پیوستگی
$A \succ B \succ C \;\Rightarrow\; \exists p \;\; [p, A;\; 1-p, C] \sim B$	
Substitutability	جانشانی‌پذیری
$A \sim B \;\Rightarrow\; [p, A;\; 1-p, C] \sim [p, B; 1-p, C]$	
Monotonicity	یکنوایی
$A \succ B \;\Rightarrow\; (p \geq q \;\Leftrightarrow\; [p, A;\; 1-p, B] \succsim [q, A;\; 1-q, B])$	

نقض قیدهای فوق، منجر به عدم رسیونالیته‌ی بدیهی می‌شود.

نظریه‌ی سودمندی

ترجیح‌های رسیونال: نتیجه‌ی نقض قیدها: مثال

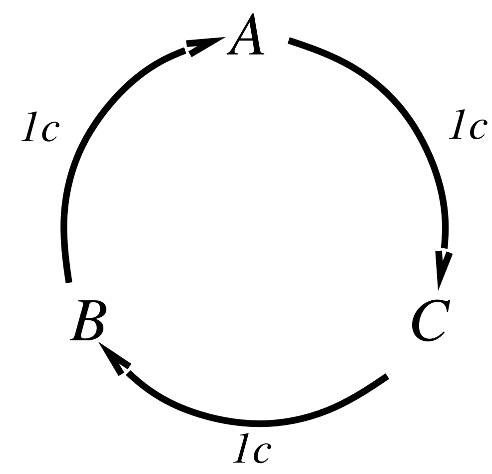
RATIONAL PREFERENCES

مثال: عاملی که ترجیح‌های غیرتراگذری دارد، می‌تواند وادار شود که همه‌ی پولش را دور بریزد!

If $B \succ C$, then an agent who has C would pay (say) 1 cent to get B

If $A \succ B$, then an agent who has B would pay (say) 1 cent to get A

If $C \succ A$, then an agent who has A would pay (say) 1 cent to get C



نظریه‌ی سودمندی

قضیه‌ی وجود تابع سودمندی برای ترجیح‌های رسیونال

قضیه

(Ramsey, 1931; von Neumann and Morgenstern, 1944)

با داشتن ترجیح‌هایی که قیدهای رسیونالیته را ارضا می‌کنند،
تابع حقیقی–مقدار U وجود دارد که

$$U(A) \geq U(B) \quad \Leftrightarrow \quad A \succsim B$$
$$U([p_1, S_1;\ \ldots\ ;\ p_n, S_n]) = \textstyle\sum_i\ p_i U(S_i)$$

اتخاذ تصمیم‌های ساده

۳

تابع‌های سودمندی

مقادیر سودمندی

UTILITIES

مقادیر سودمندی، حالت‌ها را به اعداد حقیقی نگاشت می‌دهند: **کدام اعداد؟**

$$U : S^* \longrightarrow \mathbb{R}$$

$$s \longmapsto U(s)$$

مقادیر سودمندی

روی‌کرد استاندارد برای سنجش سودمندی‌های بشری

STANDARD APPROACH TO ASSESSMENT OF HUMAN UTILITIES

حالت داده شده‌ی A را با یک لاتاری استاندارد L_p مقایسه کنید که دارای مشخصه‌ی زیر است:

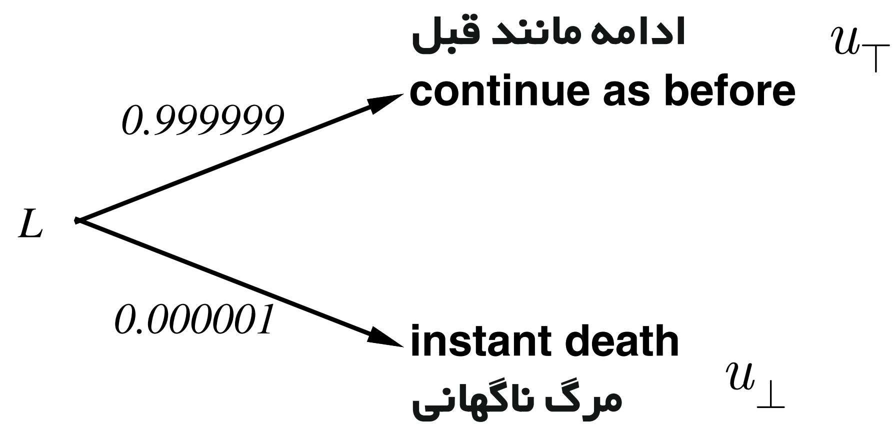
بهترین جایزه‌ی ممکن *Best Possible Prize*	$u_\top$	با احتمال p
بدترین فاجعه‌ی ممکن *Worst Possible Catastrophe*	$u_\bot$	با احتمال $1-p$

$$L = [p, u_\top; 1-p, u_\bot]$$

احتمال لاتاری p را تنظیم کنید تا

$$A \sim L_p$$

پرداخت
pay $30 ~



مقیاس‌های سودمندی

Utility scales

سودمندی نرمال‌شده *Normalized Utilitie*	میکرومورت *Micromort*	کالی *QALY: quality adjusted life year*
$u_\top = 1.0,\ u_\bot = 0.0$	شانس مرگ یک در میلیون	تعداد سال‌های زندگی با کیفیت (یک سال در سلامتی کامل بدون هیچ ناتوانی)
	کاربرد: پرداخت برای کاهش ریسک‌های تولید، ...	کاربرد: تصمیم‌گیری‌های پزشکی شامل ریسک اساسی

مقیاس‌های سودمندی

تبدیل روی تابع سودمندی

Utility scales

رفتار عامل با اعمال یک تبدیل خطی مثبت روی تابع سودمندی، **بدون‌تغییر** باقی می‌ماند.

$$U'(x) = k_1 U(x) + k_2 \quad \text{where } k_1 > 0$$

در صورتی که فقط جایزه‌های قطعی موجود باشند (نبود گزینه‌های لاتاری)،
فقط سودمندی ترتیبی (ordinal utility) قابل تعیین است $\equiv$ ترتیب کلی روی جایزه‌ها

$$\Downarrow$$

رفتار عامل با اعمال هر تبدیل یکنوای صعودی روی تابع سودمندی، **بدون‌تغییر** باقی می‌ماند.

سودمندی پول

Utility of Money

پول به‌صورت یک تابع سودمندی رفتار نمی‌کند.

در یک لاتاری L با «امید ارزش پولی» $EMV(L)$، معمولاً
$$U(L) < U(EMV(L))$$
یعنی مردم ریسک‌گریز (risk-averse) هستند.

سودمندی پول متناسب با لگاریتم مقدار پول است.

سودمندی پول

مثال

Utility of Money

در یک مسابقه‌ی تلویزیونی، بر دیگر رقبا پیروز شده‌اید.
مجری مسابقه به شما پیشنهاد انتخاب می‌دهد:

- می‌توانید جایزه‌ی 1000000$ (یک میلیون دلاری) را بگیرید.
- می‌توانید با پرتاب سکه شرط‌بندی کنید (لاتاری):
  - اگر سکه رو بیاید: هیچ جایزه‌ای دریافت نمی‌کنید.
  - اگر سکه پشت بیاید: جایزه‌ی 3000000$ (سه میلیون دلاری) را می‌گیرید.

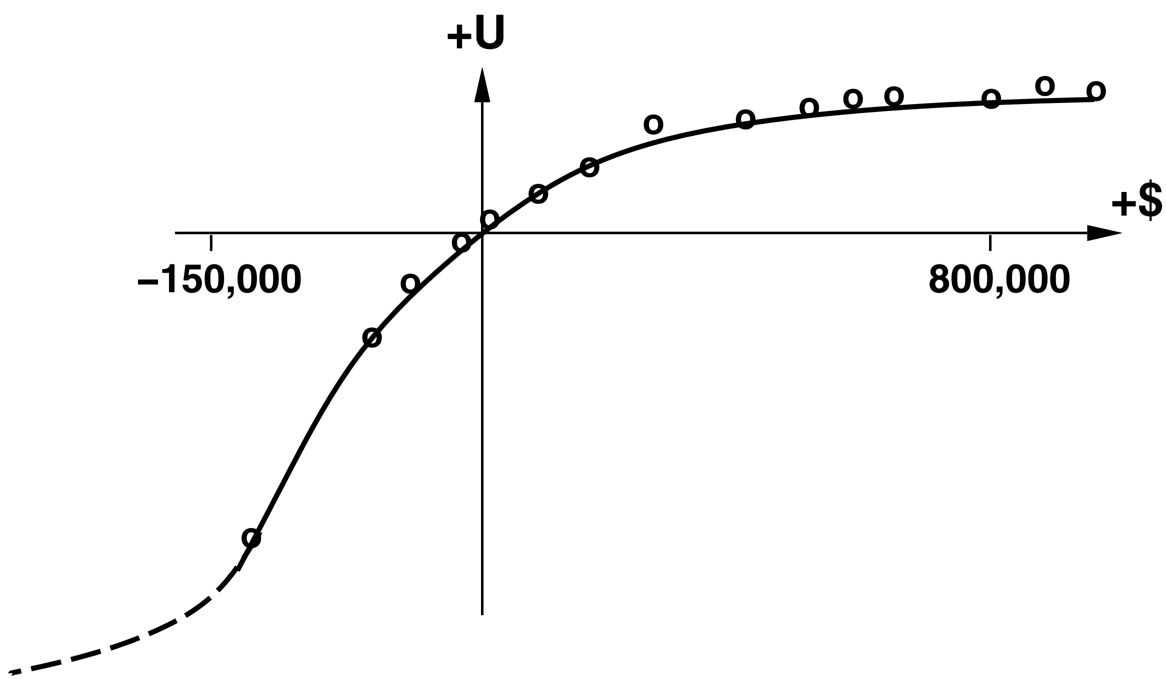
بیشتر مردم لاتاری را رد می‌کنند و همان جایزه‌ی 1000000$ را می‌گیرند
(زیرا اغلب مردم ریسک‌گریز هستند.)

سودمندی پول

منحنی سودمندی

Utility Curve

برای چه مقداری از احتمال p تفاوتی بین
جایزه‌ی x و لاتاری $[p, \$M; (1-p), \$0]$ (برای M بزرگ)
برای من وجود ندارد؟



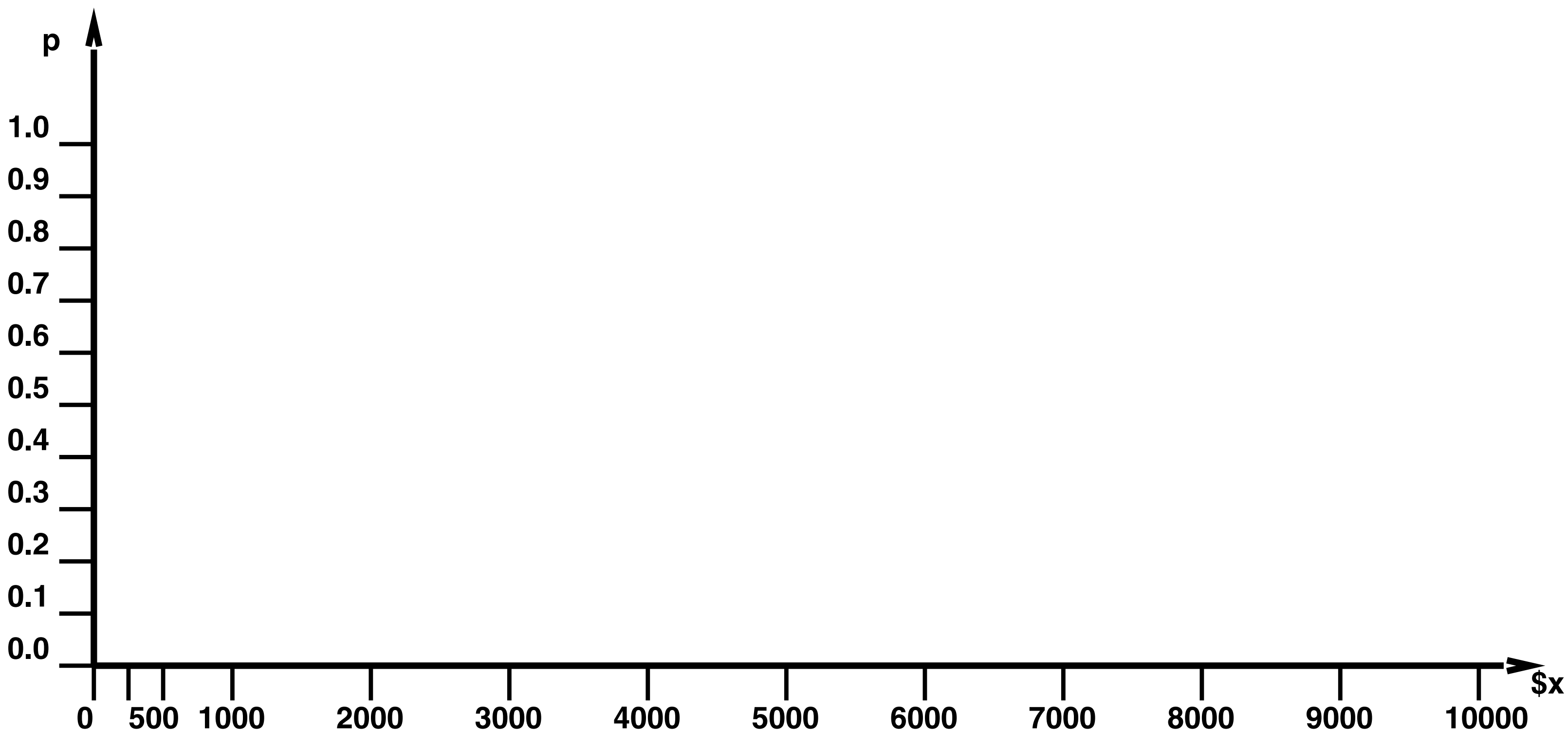
داده‌های تجربی نوعی؛ برون‌یابی با رفتار ریسک‌پذیر (risk-prone)

سودمندی پول

منحنی سودمندی برای گروه دانشجویان کلاس

Utility Curve

برای هر x مقدار p را به‌گونه‌ای تعیین کنید که **حداقل نیمی از کلاس**
لاتاری $[p, \$M; (1-p), \$0]$ (برای $M = 10000$ بزرگ) را به جایزه‌ی x ترجیح بدهند.



۴

تابع‌های سودمندی چندخصیصه‌ای

توابع سودمندی چندخصیصه‌ای

MULTIATTRIBUTE UTILITY FUNCTION

How can we handle utility functions of many variables $X_1 \ldots X_n$?
E.g., what is $U(Deaths, Noise, Cost)$?

How can complex utility functions be assessed from preference behaviour?

Idea 1: identify conditions under which decisions can be made without complete identification of $U(x_1, \ldots, x_n)$

Idea 2: identify various types of **independence** in preferences and derive consequent canonical forms for $U(x_1, \ldots, x_n)$

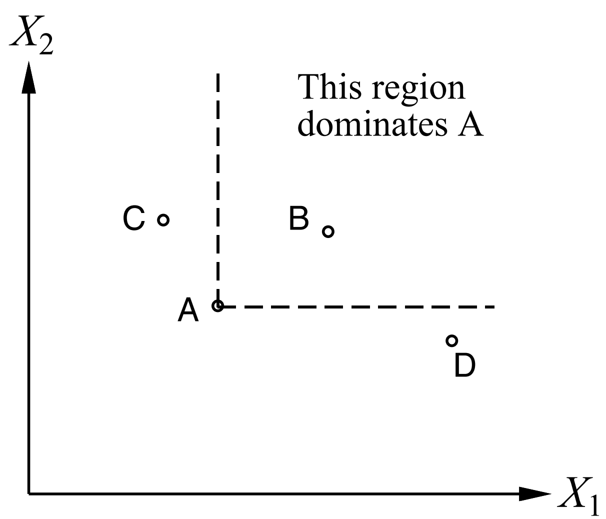
توابع سودمندی چندخصیصه‌ای

غلبه‌ی اکید

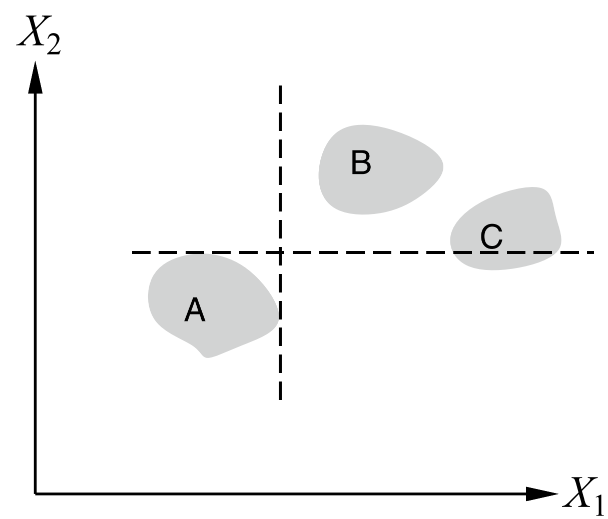
STRICT DOMINANCE

Typically define attributes such that U is monotonic in each

Strict dominance: choice B strictly dominates choice A iff
$\forall i \;\; X_i(B) \geq X_i(A)$ (and hence $U(B) \geq U(A)$)



(a)



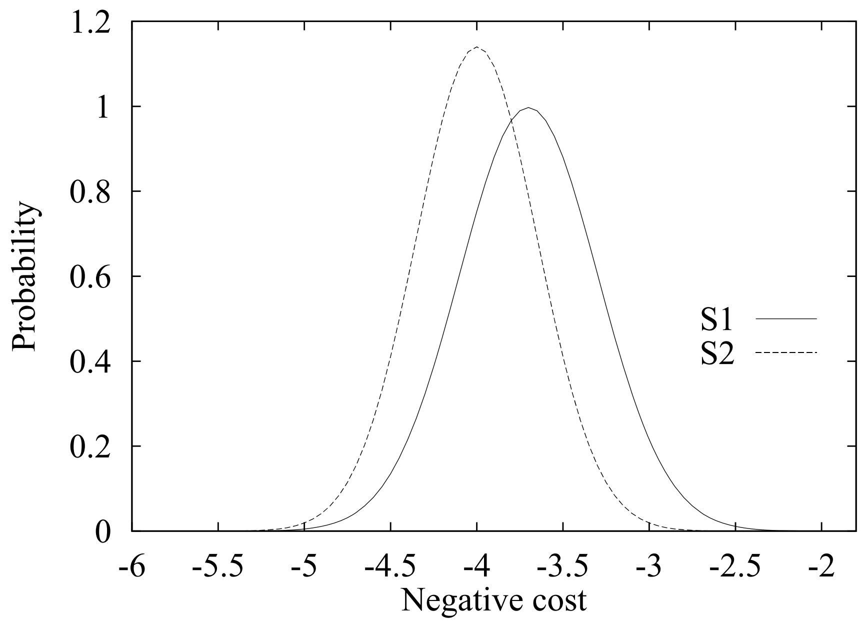
(b)

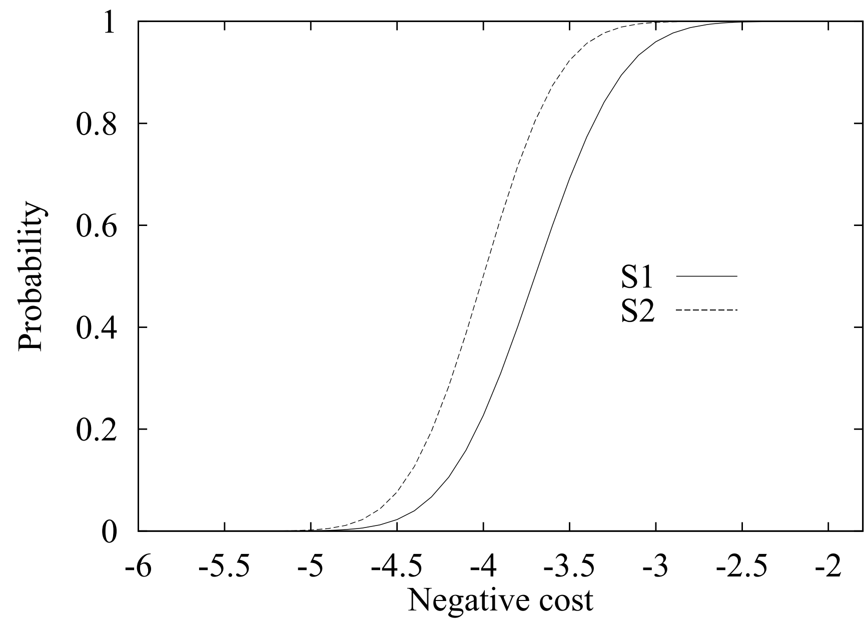
Strict dominance seldom holds in practice

توابع سودمندی چندخصیصه‌ای

غلبه‌ی اتفاقی

Strict Dominance





Distribution p_1 stochastically dominates distribution p_2 iff

$$\forall\, t \quad \int_{-\infty}^{t} p_1(x)dx \leq \int_{-\infty}^{t} p_2(t)dt$$

If U is monotonic in x, then A_1 with outcome distribution p_1 stochastically dominates A_2 with outcome distribution p_2:

$$\int_{-\infty}^{\infty} p_1(x)U(x)dx \geq \int_{-\infty}^{\infty} p_2(x)U(x)dx$$

Multiattribute case: stochastic dominance on all attributes $\Rightarrow$ optimal

توابع سودمندی چندخصیصه‌ای

Strict Dominance

Stochastic dominance can often be determined without exact distributions using **qualitative** reasoning

E.g., construction cost increases with distance from city
S_1 is closer to the city than S_2
$\Rightarrow$ S_1 stochastically dominates S_2 on cost

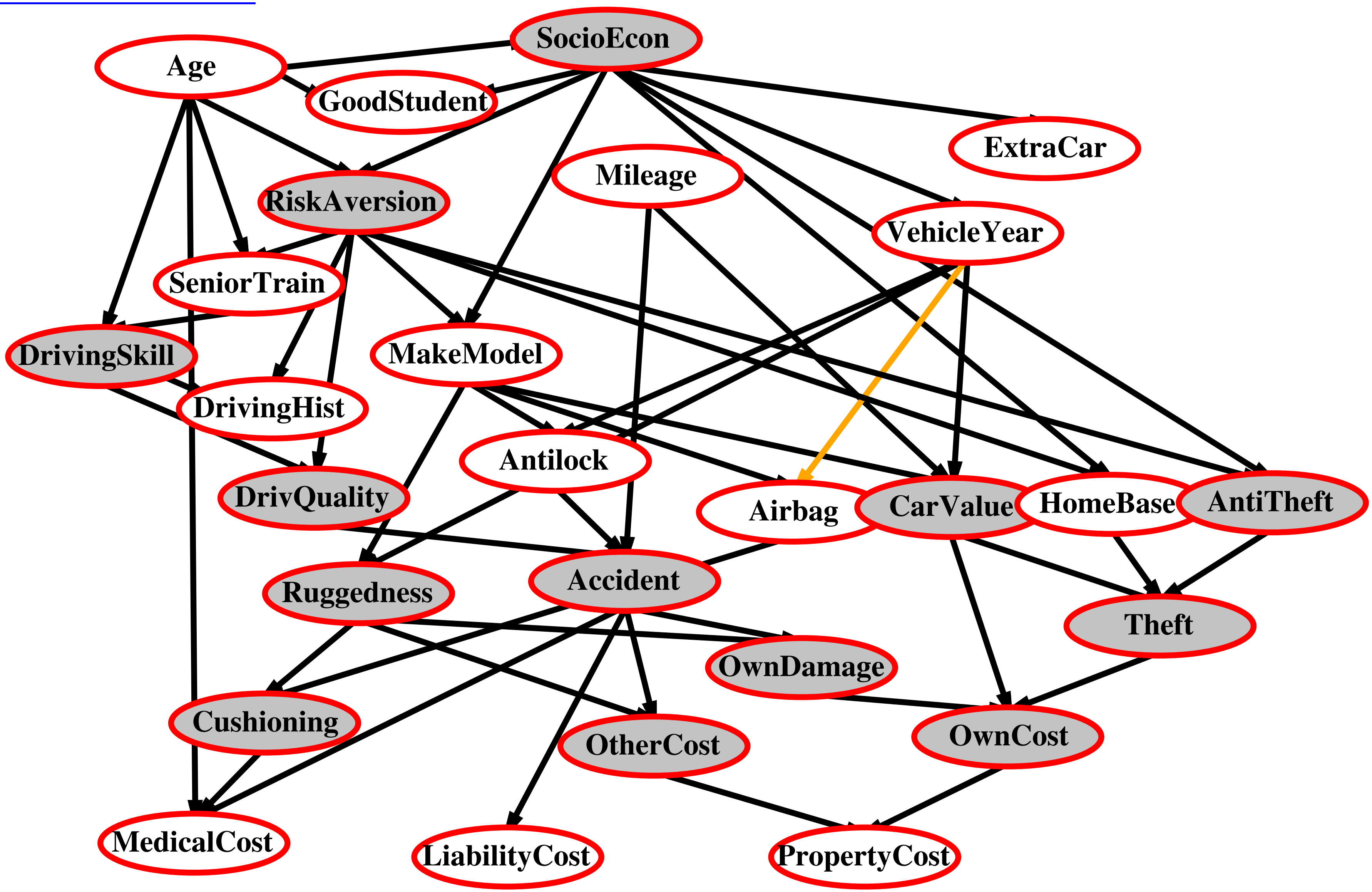
E.g., injury increases with collision speed

Can annotate belief networks with stochastic dominance information:
$X \xrightarrow{+} Y$ (X positively influences Y) means that
For every value $\mathbf{z}$ of Y's other parents $\mathbf{Z}$
$\forall\, x_1, x_2 \;\; x_1 \geq x_2 \Rightarrow \mathbf{P}(Y|x_1, \mathbf{z})$ stochastically dominates $\mathbf{P}(Y|x_2, \mathbf{z})$

توابع سودمندی چندخصیصه‌ای

غلبه‌ی اتفاقی: مشخص‌سازی: مثال (۱ از ۶)

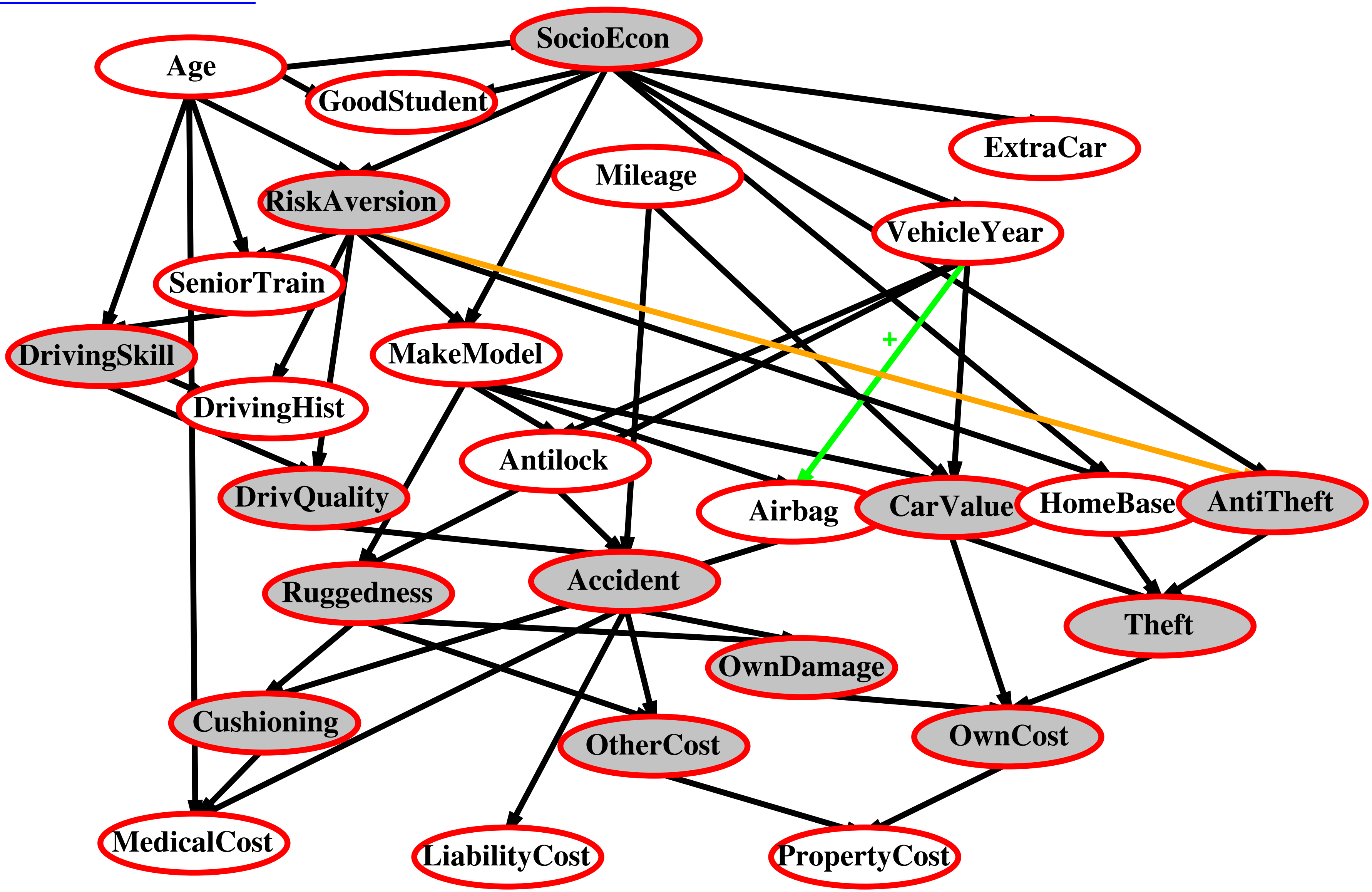
STRICT DOMINANCE



توابع سودمندی چندخصیصه‌ای

غلبه‌ی اتفاقی: مشخص‌سازی: مثال (۲ از ۶)

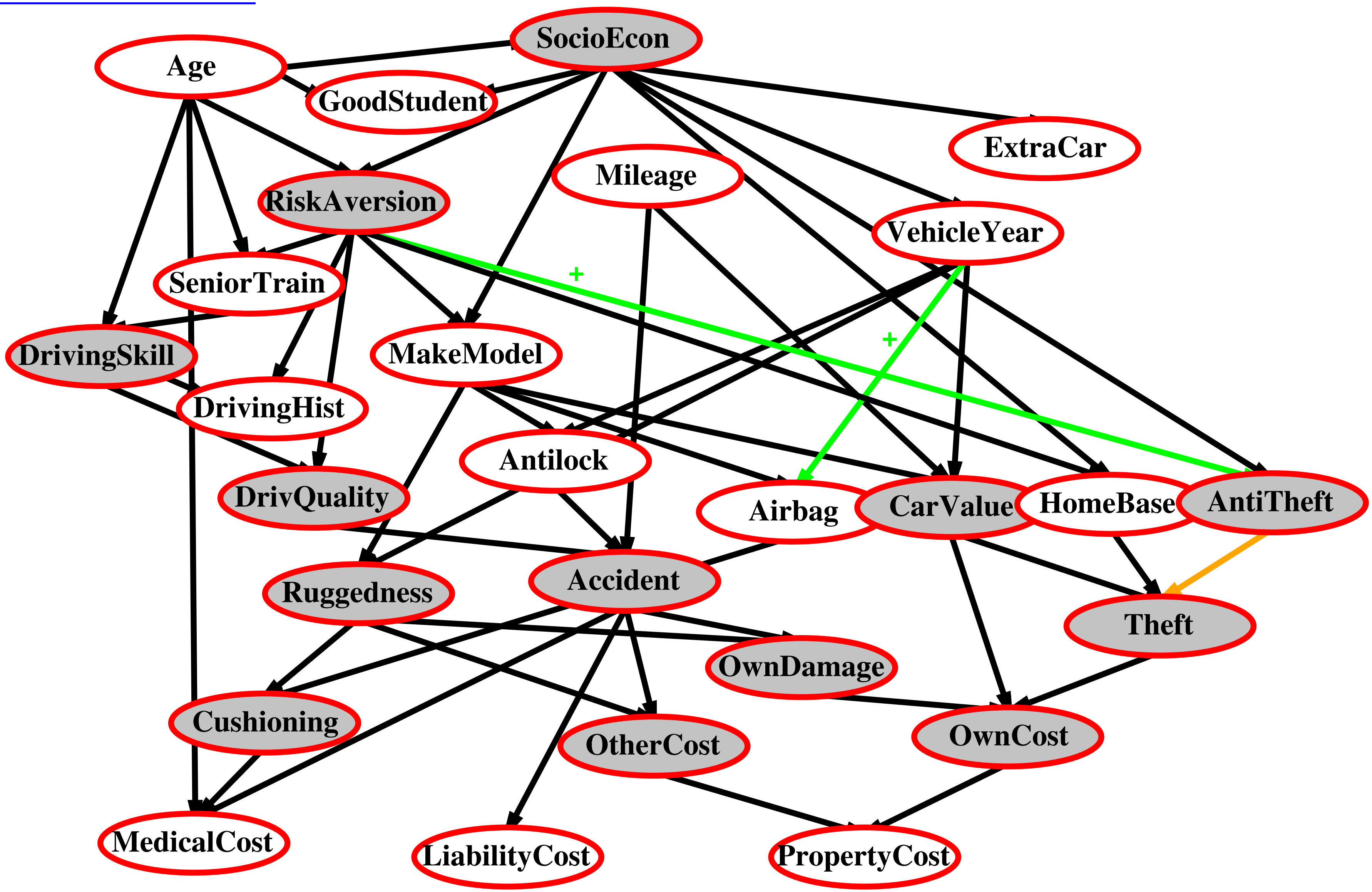
STRICT DOMINANCE



توابع سودمندی چندخصیصه‌ای

غلبه‌ی اتفاقی: مشخص‌سازی: مثال (۳ از ۶)

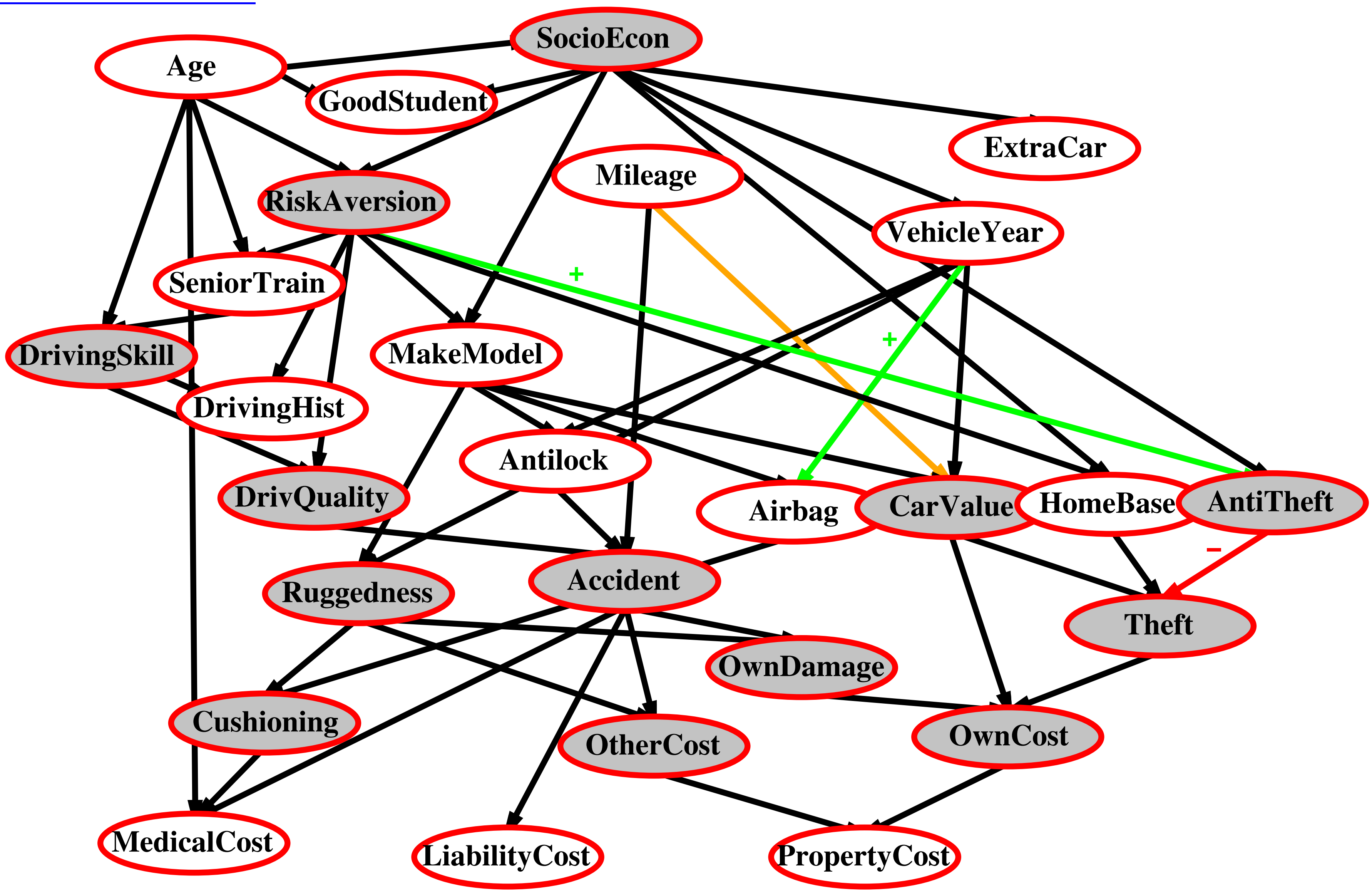
STRICT DOMINANCE



توابع سودمندی چندخصیصه‌ای

غلبه‌ی اتفاقی: مشخص‌سازی: مثال (۴ از ۶)

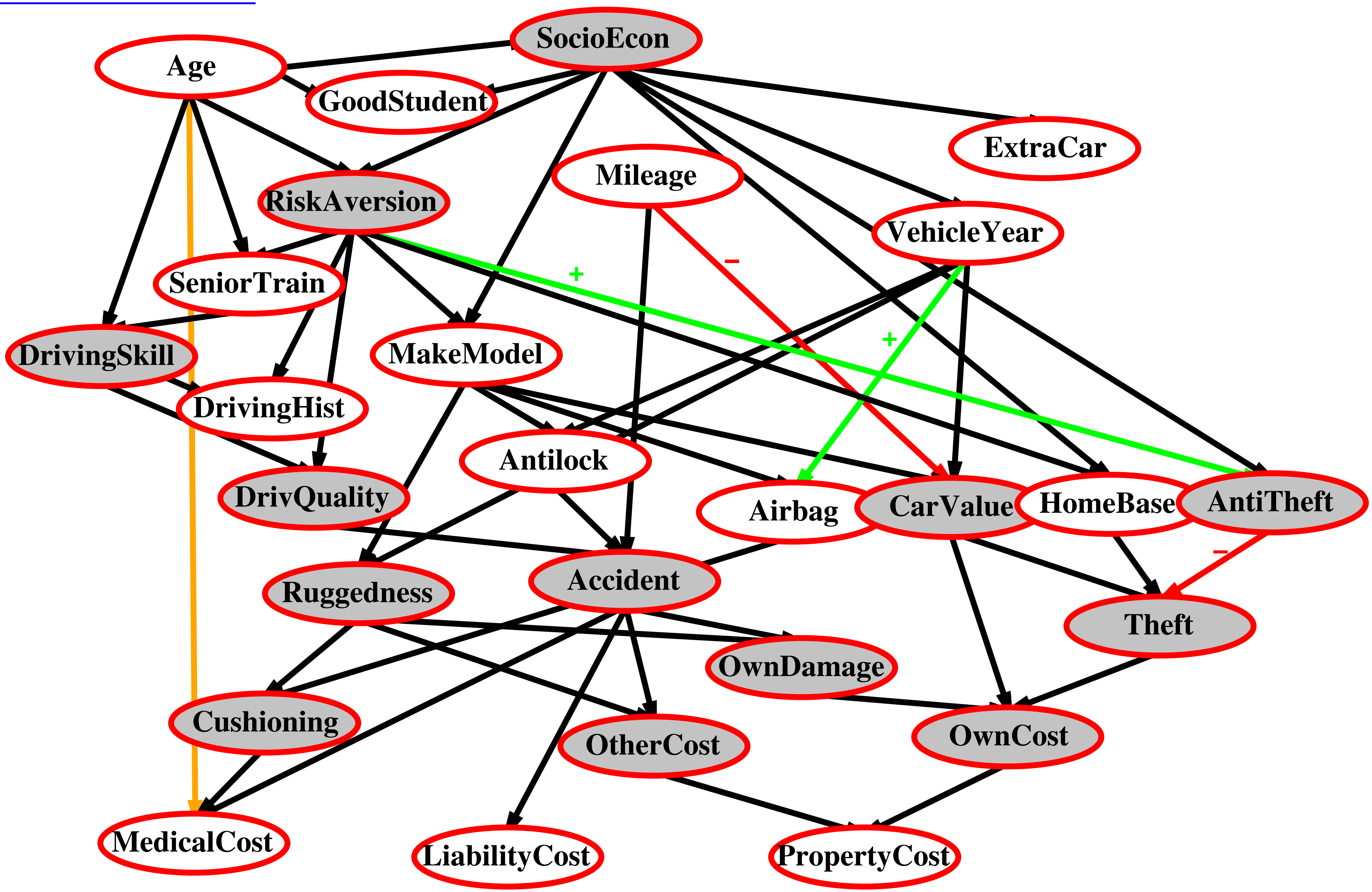
STRICT DOMINANCE



توابع سودمندی چندخصیصه‌ای

غلبه‌ی اتفاقی: مشخص‌سازی: مثال (۵ از ۶)

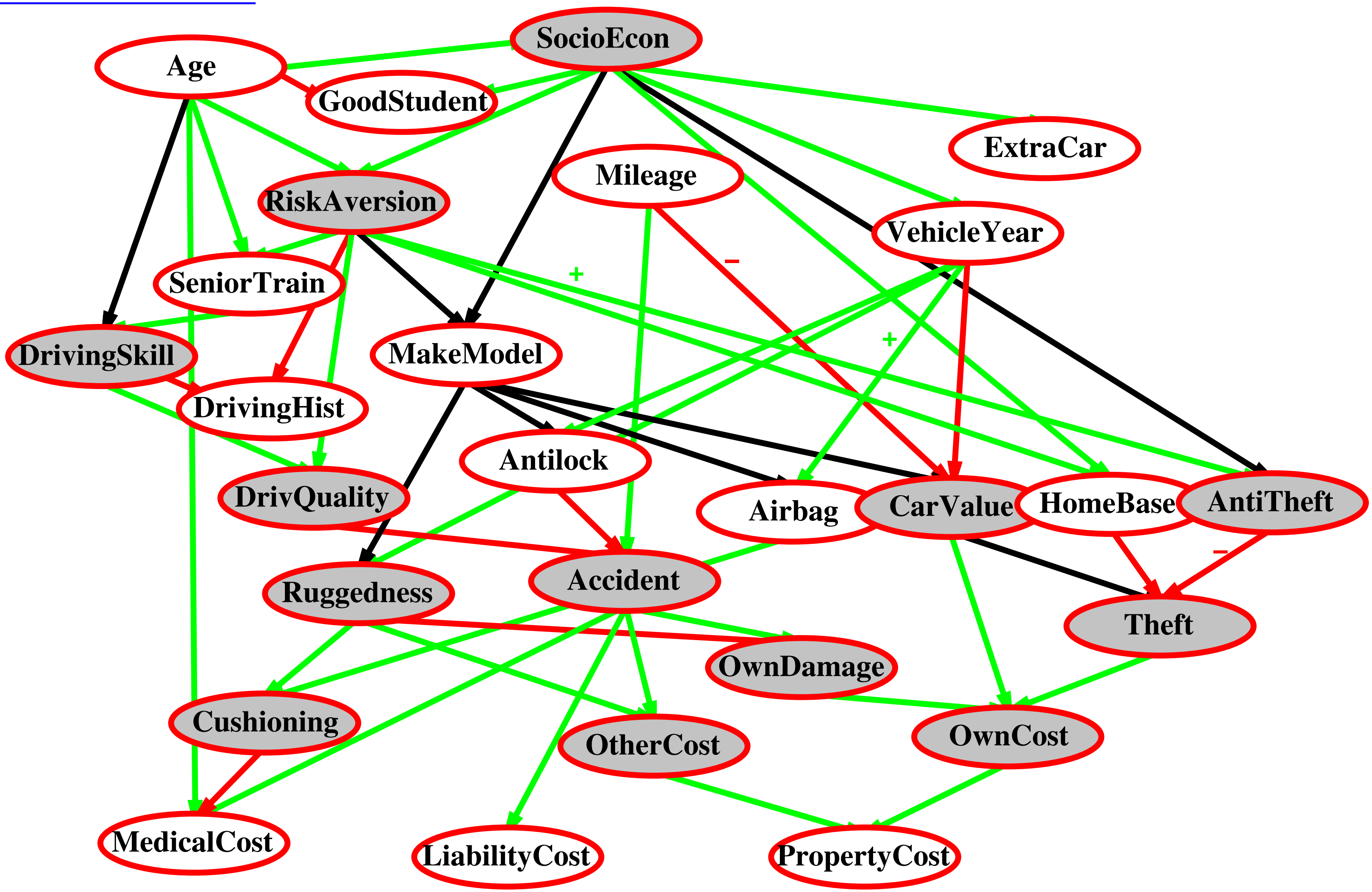
STRICT DOMINANCE



توابع سودمندی چندخصیصه‌ای

غلبه‌ی اتفاقی: مشخص‌سازی: مثال (۶ از ۶)

STRICT DOMINANCE



ساختار ترجیح‌ها

ترجیح‌های قطعی

PREFERENCE STRUCTURE: DETERMINISTIC

X_1 and X_2 preferentially independent of X_3 iff
preference between $\langle x_1, x_2, x_3 \rangle$ and $\langle x'_1, x'_2, x_3 \rangle$
does not depend on x_3

E.g., $\langle Noise, Cost, Safety \rangle$:
$\langle$20,000 suffer, \$4.6 billion, 0.06 deaths/mpm$\rangle$ vs.
$\langle$70,000 suffer, \$4.2 billion, 0.06 deaths/mpm$\rangle$

Theorem (Leontief, 1947): if every pair of attributes is P.I. of its complement, then every subset of attributes is P.I of its complement: mutual P.I..

Theorem (Debreu, 1960): mutual P.I. $\Rightarrow$ $\exists$ additive value function:

$$V(S) = \Sigma_i V_i(X_i(S))$$

Hence assess n single-attribute functions; often a good approximation

Preference Structure: Stochastic

Need to consider preferences over lotteries:
$\mathbf{X}$ is utility-independent of $\mathbf{Y}$ iff
preferences over lotteries in $\mathbf{X}$ do not depend on $\mathbf{y}$

Mutual U.I.: each subset is U.I of its complement
$\Rightarrow$ $\exists$ multiplicative utility function:

$$U = k_1U_1 + k_2U_2 + k_3U_3 + k_1k_2U_1U_2 + k_2k_3U_2U_3 + k_3k_1U_3U_1 + k_1k_2k_3U_1U_2U_3$$

Routine procedures and software packages for generating preference tests to identify various canonical families of utility functions

اتخاذ تصمیم‌های ساده

۵

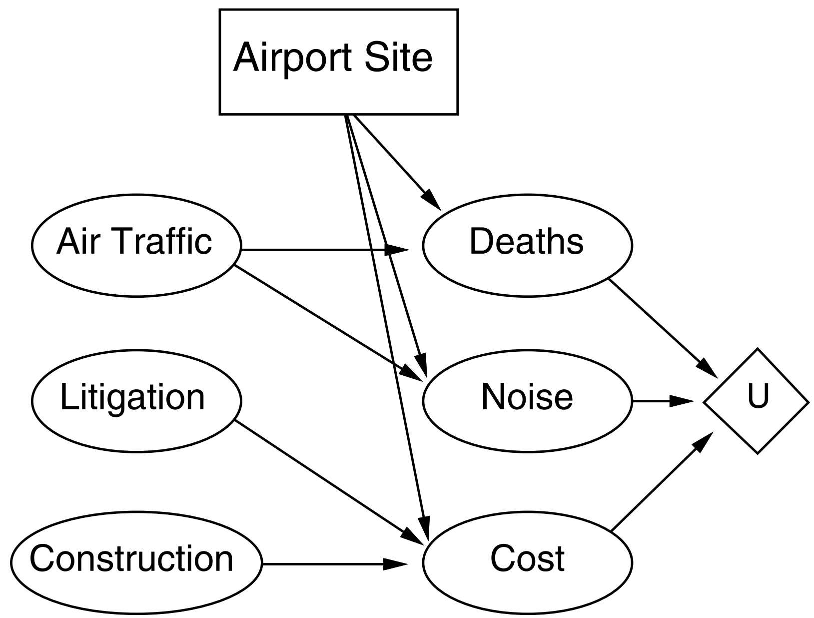
شبکه‌های تصمیم

شبکه‌های تصمیم

DECISION NETWORKS

با اضافه کردن گره‌های کنش و گره‌های سودمندی به شبکه‌های بیزی

برای ایجاد امکان تصمیم‌گیری رسیونال



الگوریتم:

برای هر مقدار از گره‌ی کنش

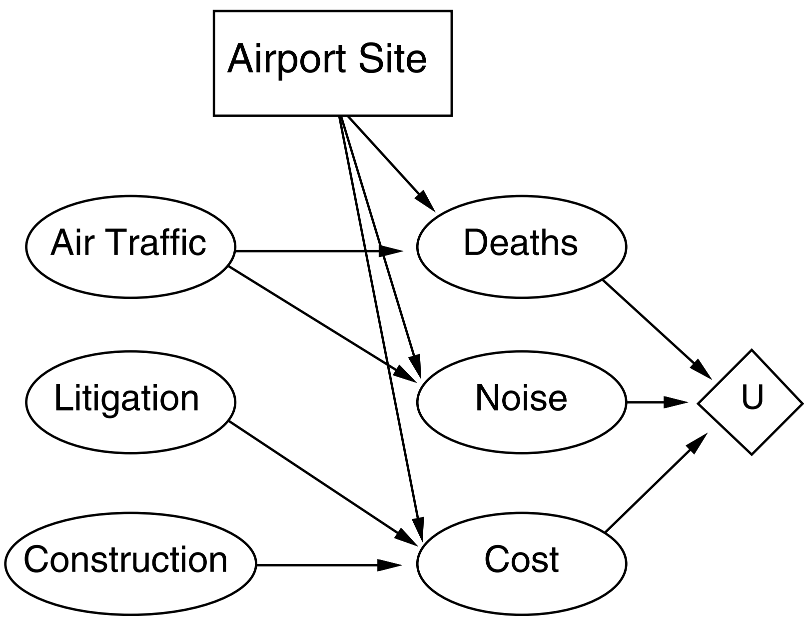
مقدار امید گره‌ی سودمندی را به شرط داشتن کنش و شاهد محاسبه کنید.

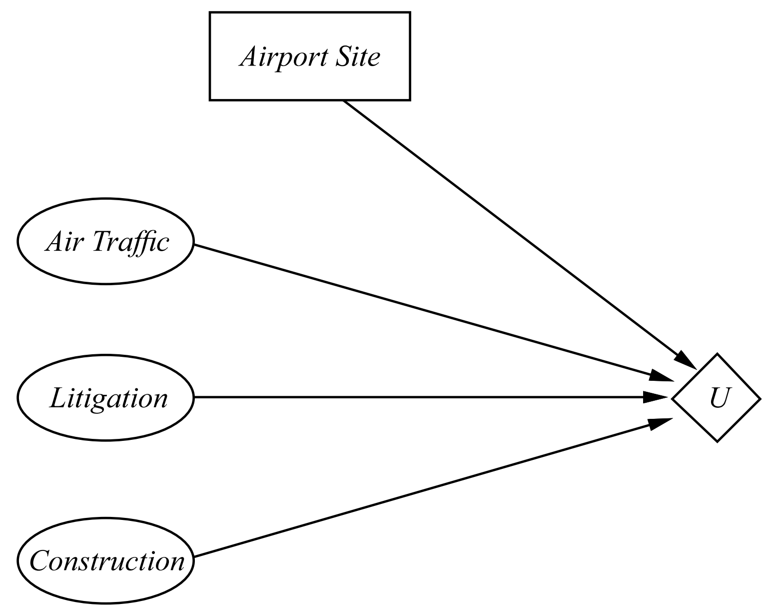
کنش MEU را برگردانید.

شبکه‌های تصمیم

[Decision Networks]

با اضافه کردن گره‌های کنش و گره‌های سودمندی به شبکه‌های بیزی
برای ایجاد امکان تصمیم‌گیری رسیونال





با کنار گذاشتن گره‌های تصادفی

هوش مصنوعی

اتخاذ تصمیم‌های ساده

۶

ارزش اطلاعات

Value of information

Idea: compute value of acquiring each possible piece of evidence
Can be done **directly from decision network**

Example: buying oil drilling rights
 Two blocks A and B, exactly one has oil, worth k
 Prior probabilities 0.5 each, mutually exclusive
 Current price of each block is $k/2$
 "Consultant" offers accurate survey of A. Fair price?

Solution: compute expected value of information
 = expected value of best action given the information
  minus expected value of best action without information
Survey may say "oil in A" or "no oil in A", **prob. 0.5 each** (given!)
 = [$0.5 \times$ value of "buy A" given "oil in A"
  + $0.5 \times$ value of "buy B" given "no oil in A"]
  – 0
 $= (0.5 \times k/2) + (0.5 \times k/2) - 0 = k/2$

ارزش اطلاعات

فرمول عمومی

VALUE OF INFORMATION

Current evidence E, current best action α
Possible action outcomes S_i, potential new evidence E_j

$$EU(\alpha|E) = \max_a \Sigma_i\ U(S_i)\ P(S_i|E,a)$$

Suppose we knew $E_j = e_{jk}$, then we would choose $\alpha_{e_{jk}}$ s.t.

$$EU(\alpha_{e_{jk}}|E, E_j = e_{jk}) = \max_a \Sigma_i\ U(S_i)\ P(S_i|E,a,E_j = e_{jk})$$

E_j is a random variable whose value is *currently* unknown
$\Rightarrow$ must compute expected gain over all possible values:

$$VPI_E(E_j) = \left(\Sigma_k\ P(E_j = e_{jk}|E) EU(\alpha_{e_{jk}}|E, E_j = e_{jk})\right) - EU(\alpha|E)$$

(VPI = value of perfect information)

VALUE OF INFORMATION

Nonnegative—in **expectation**, not **post hoc**

$$\forall j, E \;\; VPI_E(E_j) \geq 0$$

Nonadditive—consider, e.g., obtaining E_j twice

$$VPI_E(E_j, E_k) \neq VPI_E(E_j) + VPI_E(E_k)$$

Order-independent

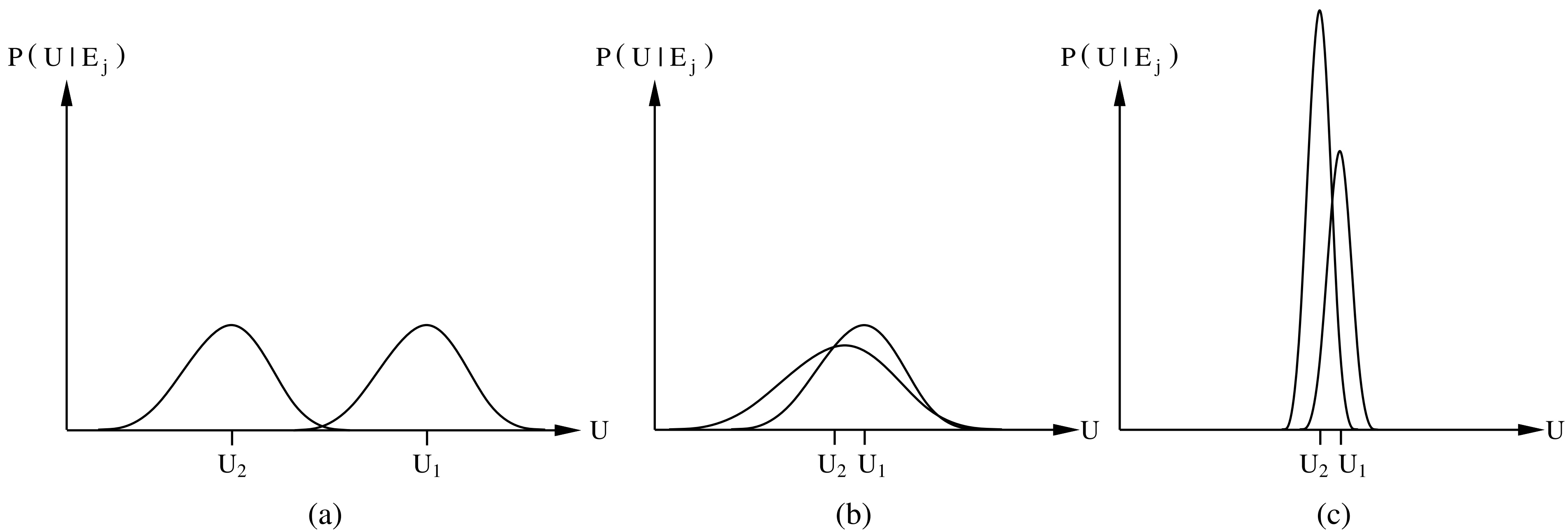
$$VPI_E(E_j, E_k) = VPI_E(E_j) + VPI_{E,E_j}(E_k) = VPI_E(E_k) + VPI_{E,E_k}(E_j)$$

Note: when more than one piece of evidence can be gathered,
maximizing VPI for each to select one is not always optimal
$\Rightarrow$ evidence-gathering becomes a **sequential** decision problem

ارزش اطلاعات

رفتارهای کیفی

QUALITATIVE BEHAVIORS



a) Choice is obvious, information worth little
b) Choice is nonobvious, information worth a lot
c) Choice is nonobvious, information worth little

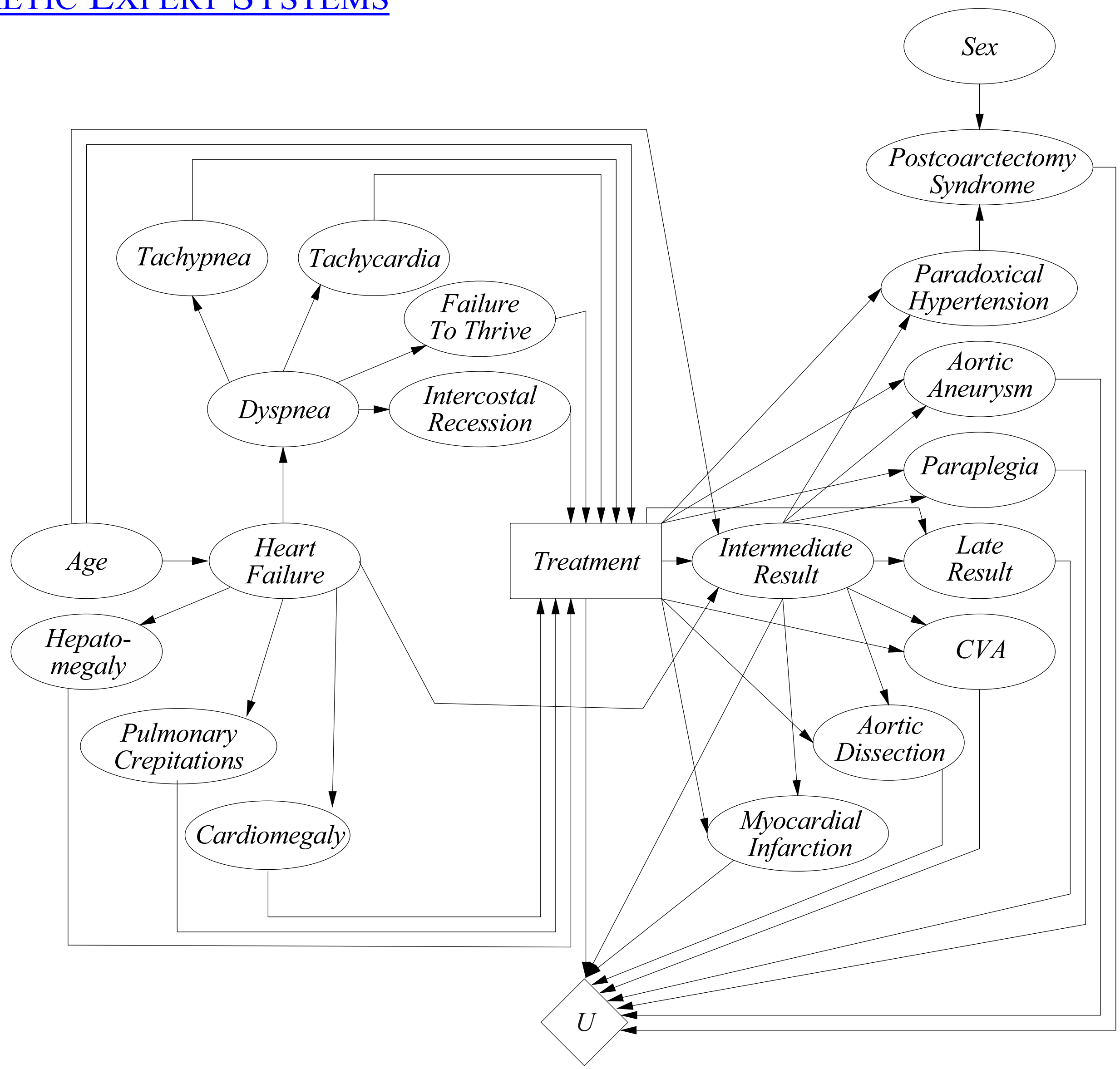
هوش مصنوعی

اتخاذ تصمیم‌های ساده

۷

سیستم‌های خبره‌ی نظریه تصمیمی

DECISION-THEORETIC EXPERT SYSTEMS



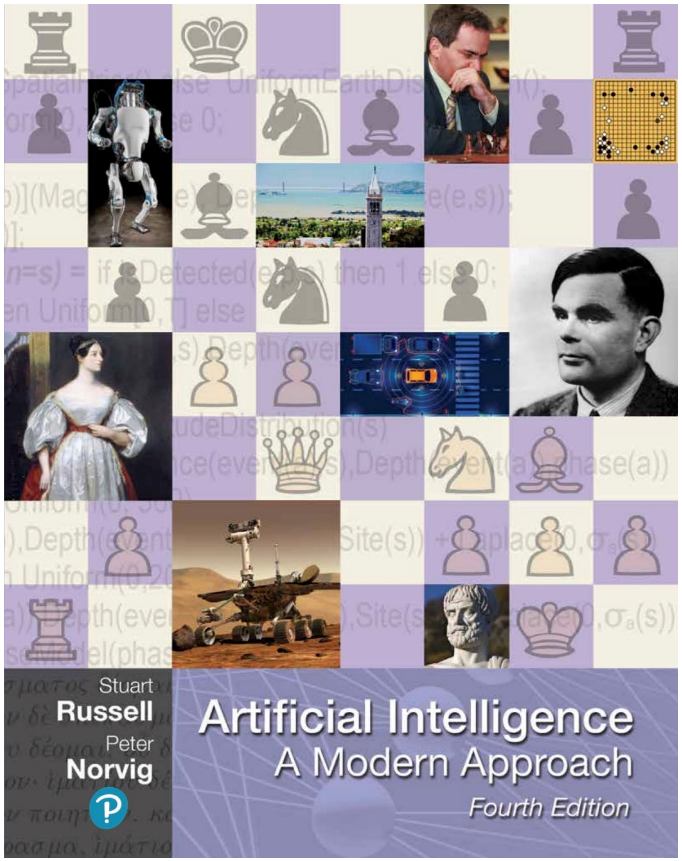
هوش مصنوعی

اتخاذ تصمیم‌های ساده

۸

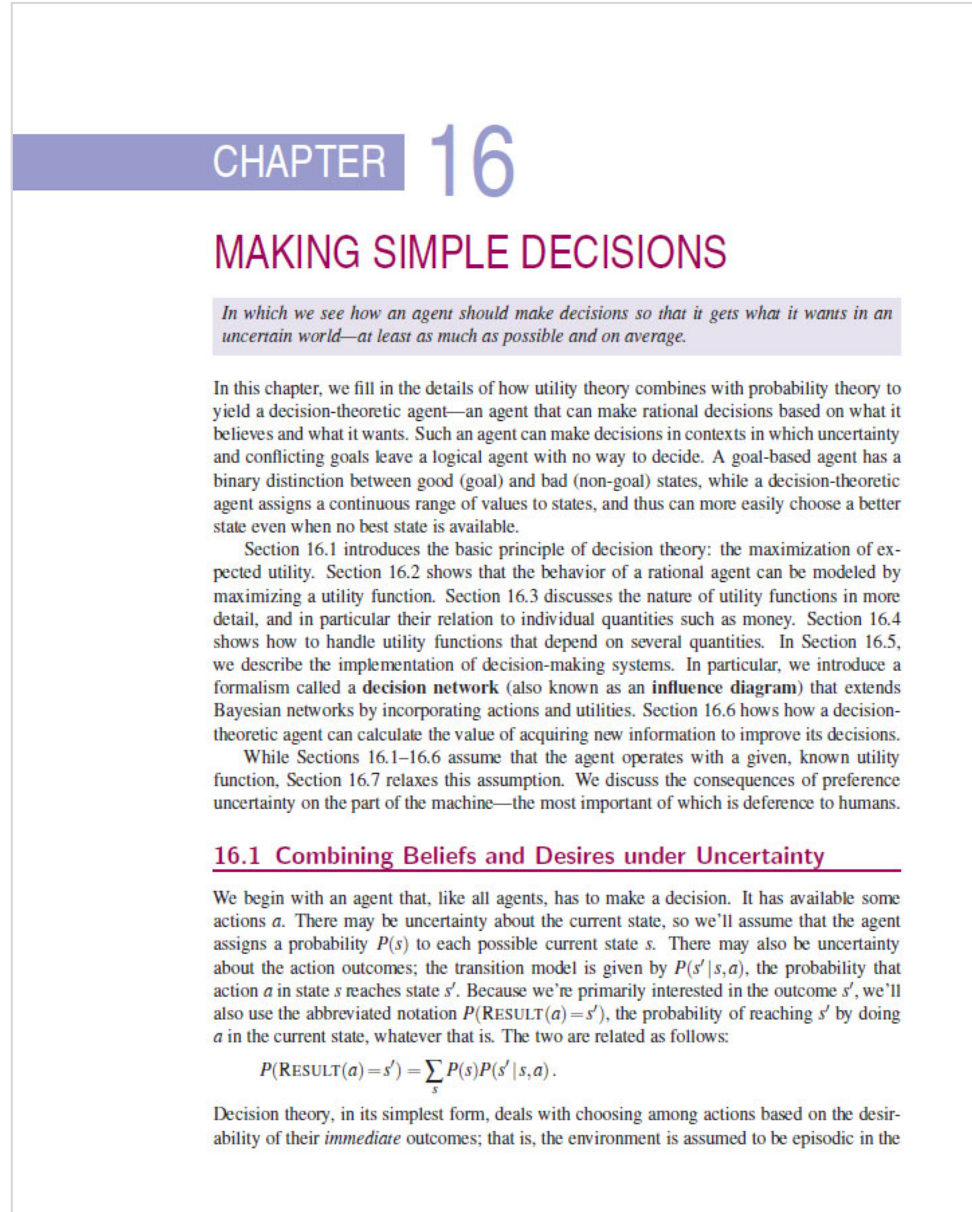
منابع، مطالعه، تکلیف

منبع اصلی



Stuart Russell and Peter Norvig,
Artificial Intelligence: A Modern Approach,
4th Edition, Prentice Hall, 2020.

Chapter 16



CHAPTER 16

MAKING SIMPLE DECISIONS

In which we see how an agent should make decisions so that it gets what it wants in an uncertain world—at least as much as possible and on average.

In this chapter, we fill in the details of how utility theory combines with probability theory to yield a decision-theoretic agent—an agent that can make rational decisions based on what it believes and what it wants. Such an agent can make decisions in contexts in which uncertainty and conflicting goals leave a logical agent with no way to decide. A goal-based agent has a binary distinction between good (goal) and bad (non-goal) states, while a decision-theoretic agent assigns a continuous range of values to states, and thus can more easily choose a better state even when no best state is available.

Section 16.1 introduces the basic principle of decision theory: the maximization of expected utility. Section 16.2 shows that the behavior of a rational agent can be modeled by maximizing a utility function. Section 16.3 discusses the nature of utility functions in more detail, and in particular their relation to individual quantities such as money. Section 16.4 shows how to handle utility functions that depend on several quantities. In Section 16.5, we describe the implementation of decision-making systems. In particular, we introduce a formalism called a **decision network** (also known as an **influence diagram**) that extends Bayesian networks by incorporating actions and utilities. Section 16.6 hows how a decision-theoretic agent can calculate the value of acquiring new information to improve its decisions.

While Sections 16.1–16.6 assume that the agent operates with a given, known utility function, Section 16.7 relaxes this assumption. We discuss the consequences of preference uncertainty on the part of the machine—the most important of which is deference to humans.

16.1 Combining Beliefs and Desires under Uncertainty

We begin with an agent that, like all agents, has to make a decision. It has available some actions a. There may be uncertainty about the current state, so we'll assume that the agent assigns a probability $P(s)$ to each possible current state s. There may also be uncertainty about the action outcomes; the transition model is given by $P(s' \mid s,a)$, the probability that action a in state s reaches state s'. Because we're primarily interested in the outcome s', we'll also use the abbreviated notation $P(\textsc{Result}(a)=s')$, the probability of reaching s' by doing a in the current state, whatever that is. The two are related as follows:

$$P(\textsc{Result}(a)=s') = \sum_s P(s)P(s' \mid s,a)\,.$$

Decision theory, in its simplest form, deals with choosing among actions based on the desirability of their *immediate* outcomes; that is, the environment is assumed to be episodic in the