بسم الله الرحمن الرحیم

هوش مصنوعی

فصل ۱۳

کمی‌سازی عدم اطمینان

Quantifying Uncertainty

کاظم فولادی
دانشکده مهندسی برق و کامپیوتر
دانشگاه تهران

http://courses.fouladi.ir/ai

هوش مصنوعی

کمی‌سازی عدم اطمینان

۱

کنش
تحت
عدم‌اطمینان

مشکلات روش‌های مبتنی بر منطق قطعی در عمل

مشکلات روش‌های مبتنی بر منطق قطعی در عمل		
تفسیر مشاهده‌پذیری جزئی	**بزرگ شدن طرح اقتضائی**	**عدم وجود تضمین در طرح**
هنگام تفسیر اطلاعات حسگری جزئی، یک عامل منطقی باید همه‌ی توضیحات منطقی ممکن برای مشاهدات را بررسی کند (بدون توجه به میزان شانس آنها) ⇓ بازنمایی‌های حالت باور بزرگ و پیچیده‌ی غیر ممکن	یک طرح اقتضائی درست که همه‌ی سرانجام‌ها را در نظر می‌گیرد، می‌تواند بسیار بزرگ شود و باید اقتضائات غیرمحتمل را نیز در نظر بگیرد.	گاهی طرحی وجود ندارد که تضمین کند به هدف می‌رسیم؛ اما عامل باید کنش کند! ⇓ باید راهی برای مقایسه‌ی مزایای طرح‌هایی که تضمین نمی‌شوند، داشته باشیم

مشکلات روش‌های مبتنی بر منطق قطعی در عمل

مثال: رسیدن به پرواز هواپیما با تاکسی خودکار

$A_t =$ کنش حرکت به سمت فرودگاه t دقیقه قبل از پرواز

آیا A_t من را به‌موقع می‌رساند؟

مشکلات			
مشاهده‌پذیری جزئی	**حسگرهای نویزی**	**عدم اطمینان از برآمد کنش‌ها**	**دشواری پیش‌بینی ترافیک**
وضعیت جاده طرح سایر راننده‌ها، …	گزارش‌های ترافیکی رادیو	مثلاً پنچر شدن تایر، …	پیچیدگی بالای مدل‌سازی و پیش‌بینی ترافیک

بنابراین، روی‌کرد منطقی خالص، منجر به یکی از این دو مورد می‌شود:

ریسک نادرست بودن	رسیدن به نتایج بسیار ضعیف برای تصمیم‌گیری
A_{25} من را به‌موقع می‌رساند.	A_{25} من را به‌موقع می‌رساند اگر هیچ تصادفی روی پل رخ نداده باشد و باران نیاید و تایرهای ماشین سالم باقی بمانند و …

می‌توان گفت قاعدتاً A_{1440} من را به‌موقع می‌رساند؛ اما باید کل شب را در فرودگاه بمانم!

مشکلات روش‌های مبتنی بر منطق قطعی در عمل

مثال: دامنه‌ی دندان‌پزشکی

تشخیص دلیل دندان‌درد یک بیمار دندان‌پزشکی

نوشتن قواعد برای **تشخیص** دندان‌پزشکی با استفاده از منطق گزاره‌ای:

$$Toothache \Rightarrow Cavity\,.$$

علت دندان‌درد کرم‌خوردگی است

مشکل این است که این قاعده غلط است!
همه‌ی بیماران دارای دندان‌درد دارای کرم‌خوردگی نیستند؛
برخی بیماری لثه دارند، برخی آبسه دارند، یا یکی از مشکلات متعدد دیگر ...

$$Toothache \Rightarrow Cavity \vee GumProblem \vee Abscess \ldots$$

متاسفانه برای درست کردن این قاعده،
باید تقریباً یک لیست نامحدود از مشکلات ممکن را اضافه کنیم!
اگر سعی کنیم این قاعده را در قالب قاعده‌ی علّی هم بنویسیم:

$$Cavity \Rightarrow Toothache\,.$$

کرم‌خوردگی موجب دندان‌درد می‌شود

باز هم نادرست است: همه‌ی کرم‌خوردگی‌ها موجب درد نمی‌شود!

تنها راه برای درست نوشتن این قاعده: باید آن را به لحاظ منطقی جامع کنیم (که عملاً نمی‌شود).

Prepared by Kazim Fouladi | Spring 2017 | 3rd Edition

مشکلات روش‌های مبتنی بر منطق قطعی در عمل

دلایل شکست روی‌کرد منطقی

دلایل شکست روی‌کرد منطقی در مسئله‌ای مانند تشخیص پزشکی		
تنبلی *Laziness*	**ناآگاهی نظری** *Theoretical Ignorance*	**ناآگاهی عملی** *Practical Ignorance*
کار زیاد لازم برای تنظیم قواعد	ناقص بودن نظریه‌های علمی	عدم امکان آزمون همه‌ی شرایط
فهرست کردن مجموعه‌ی کاملی از مقدم‌ها و تالی‌های مورد نیاز برای ایجاد قواعد بدون استثنا، کار و زمان بسیاری می‌برد و به‌کارگیری چنین قواعدی بسیاری مشکل است.	دانش پزشکی هیچ نظریه‌ی کامل و جامعی برای این زمینه ندارد!	حتی اگر تمامی قواعد را هم بشناسیم، باز هم ممکن است نتوانیم در مورد بیمار نظر قطعی ارائه بدهیم، مثلاً ممکن است نتوانیم همه‌ی آزمایش‌های لازم را انجام بدهیم.

روش‌های برخورد با عدم اطمینان

روش‌های برخورد با عدم اطمینان (استدلال نامطمئن)				
منطق غیریکنوا/پیش‌فرض *Default/Nonmonotonic Logic*	**مبتنی بر قاعده** *Rule-based*	**نظریه دمپستر-شافر** *Dempster–Shafer theory*	**منطق فازی** *Fyzzy Logic*	**نظریه احتمال** *Probability Theory*
استدلال‌های کیفی *Qualitative Reasoning*	با فاکتور فاج *Using Fudge Factor*	بازنمایی ناآگاهی *Representing Ignorance*	بازنمایی سربستگی *Representing Vagueness*	درجه‌ی باور به درستی رویداد *Belief Degree for Truth*

روش‌های برخورد با عدم اطمینان

منطق غیریکنوا / منطق پیش‌فرض

روش‌های برخورد با عدم اطمینان (استدلال نامطمئن)				
منطق غیریکنوا/پیش‌فرض *Default/Nonmonotonic Logic*	مبتنی بر قاعده *Rule-based*	نظریه دمپستر-شافر *Dempster-Shafer theory*	منطق فازی *Fuzzy Logic*	نظریه احتمال *Probability Theory*
استدلال‌های کیفی *Qualitative Reasoning*	با فاکتور قاطع *Using Fudge Factor*	بازنمایی ناآگاهی *Representing Ignorance*	بازنمایی سربستگی *Representing Vagueness*	درجه‌ی باور به درستی رویداد *Belief Degree for Truth*

استفاده از استدلال‌های کیفی (مشابه منطق انسانی) به‌جای محاسبات عددی

منطق پیش‌فرض:

برخورد با نتایج به صورت باور

تا زمانی که دلیل بهتری برای باور به چیز دیگری پیدا شود.

مثال: به‌فرض، ماشین من لاستیک پنچر ندارد،

به‌فرض، A_{25} من را به‌موقع می‌رساند مگر اینکه با شاهدی تناقض پیدا کند.

مشکلات: ❋ چه فرض‌هایی مستدل هستند؟ ❋ چگونگی با تناقض‌ها برخورد کنیم؟

روش‌های برخورد با عدم اطمینان

مبتنی بر قاعده

روش‌های برخورد با عدم اطمینان (استدلال نامطمئن)				
منطق غیریکنوا *Nonmonotonic Logic*	**مبتنی بر قاعده** *Rule-based*	نظریه دمپستر-شافر *Dempster-Shafer theory*	منطق فازی *Fuzzy Logic*	نظریه احتمال *Probability Theory*
استدلال‌های کیفی *Qualitative Reasoning*	**با فاکتور فاج** *Using Fudge Factor*	بازنمایی ناآگاهی *Representing Ignorance*	بازنمایی سربستگی *Representing Vagueness*	درجه‌ی باور به درستی رویداد *Belief Degree for Truth*

ساخت سیستم‌های مبتنی بر قاعده‌ی منطقی،
با اضافه کردن نوعی عامل فاج به هر قاعده، برای برخورد با عدم اطمینان

$$A_{25} \mapsto_{0.3} AtAirportOnTime$$
$$Sprinkler \mapsto_{0.99} WetGrass$$
$$WetGrass \mapsto_{0.7} Rain$$

مشکلات: ❋ چگونگی ترکیب نتایج: مثلاً $Sprinkler \mapsto Rain$?? با چه درجه‌ای؟

روش‌های برخورد با عدم اطمینان

نظریه دمپستر- شافر

روش‌های برخورد با عدم اطمینان (استدلال نامطمئن)				
منطق غیریکنواخت Nonmonotonic Logic	مبتنی بر قاعده Rule-based	**نظریه دمپستر- شافر** *Dempster–Shafer theory*	منطق فازی Fuzzy Logic	نظریه احتمال Probability Theory
استدلال‌های کیفی Qualitative Reasoning	با فاکتور فاج Using Fudge Factor	**بازنمایی ناآگاهی** *Representing Ignorance*	بازنمایی سربستگی Representing Vagueness	درجه‌ی باور به درستی رویداد Belief Degree for Truth

استفاده از درجه‌های باور با مقادیر بازه‌ای (کران بالا، کران پایین)
برای بازنمایی دانایی عامل در مورد احتمال یک گزاره

مشکلات: ❋ پیچیدگی بالای محاسباتی

روش‌های برخورد با عدم اطمینان

منطق فازی

روش‌های برخورد با عدم اطمینان (استدلال نامطمئن)				
منطق غیریکنواختی *Nonmonotonic Logic*	مبتنی بر قاعده *Rule-based*	نظریه دمپستر-شافر *Dempster-Shafer theory*	**منطق فازی** *Fyzzy Logic*	نظریه احتمال *Probability Theory*
استدلال‌های کیفی *Qualitative Reasoning*	با فاکتور فاج *Using Fudge Factor*	بازنمایی ناآگاهی *Representing Ignorance*	**بازنمایی سربستگی** *Representing Vagueness*	درجه‌ی باور به درستی رویداد *Belief Degree for Truth*

هستی‌شناسی منطق فازی: اجازه دادن به **سربستگی**

یک گزاره می‌تواند تا مرتبه‌ای درست باشد **(درجه‌ی درستی)**

احتمالات و منطق معمولی تعهدات هستی‌شناختی یکسانی دارند:

گزاره‌های دنیا **درست یا نادرست** هستند،

حتی اگر عامل نامطمئن باشد که کدام یک است.

✻ مناسب برای کار کردن با مفاهیمی که حد و مرز دقیقی ندارند.

روش‌های برخورد با عدم اطمینان

نظریه‌ی احتمال

روش‌های برخورد با عدم اطمینان (استدلال نامطمئن)				
منطق غیریکنوا Nonmonotonic Logic	مبتنی بر قاعده Rule-based	نظریه دمپستر-شافر Dempster-Shafer theory	منطق فازی Fuzzy Logic	**نظریه احتمال** *Probability Theory*
استدلال‌های کیفی Qualitative Reasoning	با فاکتور فاج Using Fudge Factor	بازنمایی ناآگاهی Representing Ignorance	بازنمایی سربستگی Representing Vagueness	**درجه‌ی باور به درستی رویداد** *Belief Degree for Truth*

به هر جمله یک درجه‌ی باور بین صفر و یک نسبت داده می‌شود.

احتمالات، روشی برای جمع‌بندی عدم اطمینان ناشی از تنبلی و ناآگاهی است.

مثال: A_{25} من را به‌موقع می‌رساند با احتمال 0.04

مشکلات: ❋ چه فرض‌هایی مستدل هستند؟ ❋ چگونگی با تناقض‌ها برخورد کنیم؟

انواع منطق‌ها

منطق احتمالات

	تعهدات منطق	
	تعهدات هستی‌شناختی *Ontological Commitment*	**تعهدات معرفت‌شناختی** *Epistemological Commitment*
منطق گزاره‌ای *Propositional Logic*	واقعیت‌ها *Facts*	درست / نادرست / ناشناخته *True / False / Unknown*
منطق مرتبه اول *First-Order Logic*	واقعیت‌ها، اشیا، رابطه‌ها *Facts, Objects, Relations*	درست / نادرست / ناشناخته *True / False / Unknown*
منطق زمانی *Temporal Logic*	واقعیت‌ها، اشیا، رابطه‌ها، زمان‌ها *Facts, Objects, Relations, Times*	درست / نادرست / ناشناخته *True / False / Unknown*
نظریه‌ی احتمال *Probability Theory*	واقعیت‌ها *Facts*	درجه‌ی باور بین صفر تا یک *Degree of Belief* 0…1
منطق فازی *Fuzzy Logic*	واقعیت‌هایی با درجه‌ی درستی بین ۰ و ۱ *Facts with Degree of Truth in* [0,1]	بازه‌ی معلومی از مقادیر *Known interval value*

منشأ احتمالات

فلسفه‌ی احتمال

منشأ احتمالات		
موضع فراوانی‌گرا *Frequentist Position*	**موضع ابژکتیو** *Objective Position*	**موضع سوبژکتیو** *Subjective Position*
اعداد احتمال حاصل **تجربه** است	احتمالات نمودهای واقعی جهان هستند	احتمالات روشی برای بیان باور عامل است
احتمالات برخاسته از فراوانی نسبی پدیده‌هاست.	اشیا به طور طبیعی تمایل به رفتار غیرمطمئن دارند (نه اینکه احتمالات فقط توصیفی از درجه‌ی باور مشاهده‌کننده باشد).	احتمالات، گزاره‌ها را به حالت دانایی شخصی یک عامل مرتبط می‌کند.

احتمال

مثال

احتمال بیزی (سوبژکتیو):

احتمالات، گزاره‌ها را به حالت دانایی شخصی یک عامل مرتبط می‌کند.

$$P(A_{25}|\text{no reported accidents}) = 0.06$$

مقدار احتمال می‌تواند از تجربیات گذشته در وضعیت‌های مشابه، یاد گرفته شود.

احتمالات گزاره‌ها با آمدن شاهد جدید، تغییر می‌کند:

$$P(A_{25}|\text{no reported accidents, 5 a.m.}) = 0.15$$

(قابل مقایسه با استلزام منطقی $KB \models \alpha$ ، نه درستی گزاره)

تصمیم‌گیری رسیونال تحت عدم اطمینان

مثال

فرض می‌کنیم باورهای زیر را داریم:

$$
\begin{aligned}
P(A_{25}\ \text{gets me there on time}|\ldots) &= 0.04\\
P(A_{90}\ \text{gets me there on time}|\ldots) &= 0.70\\
P(A_{120}\ \text{gets me there on time}|\ldots) &= 0.95\\
P(A_{1440}\ \text{gets me there on time}|\ldots) &= 0.9999
\end{aligned}
$$

کدام کنش باید انتخاب شود؟

تصمیم‌گیری در مورد کنش فقط به احتمال وابسته نیست؛
بلکه به ترجیح ما در مورد اهمیت پرواز، میزان معطلی در فرودگاه و ... هم وابسته است.

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 تصمیم

نظریه‌ی عمومی تصمیم‌های رسیونال

DECISION THEORY

نظریه‌ی احتمال
Probability Theory

+

نظریه‌ی سودمندی
Utility Theory

=

نظریه‌ی تصمیم
Decision Theory

یک عامل رسیونال است

اگر و فقط اگر

کنشی با بالاترین سودمندی مورد انتظار را انتخاب کند.

(متوسط آماری)

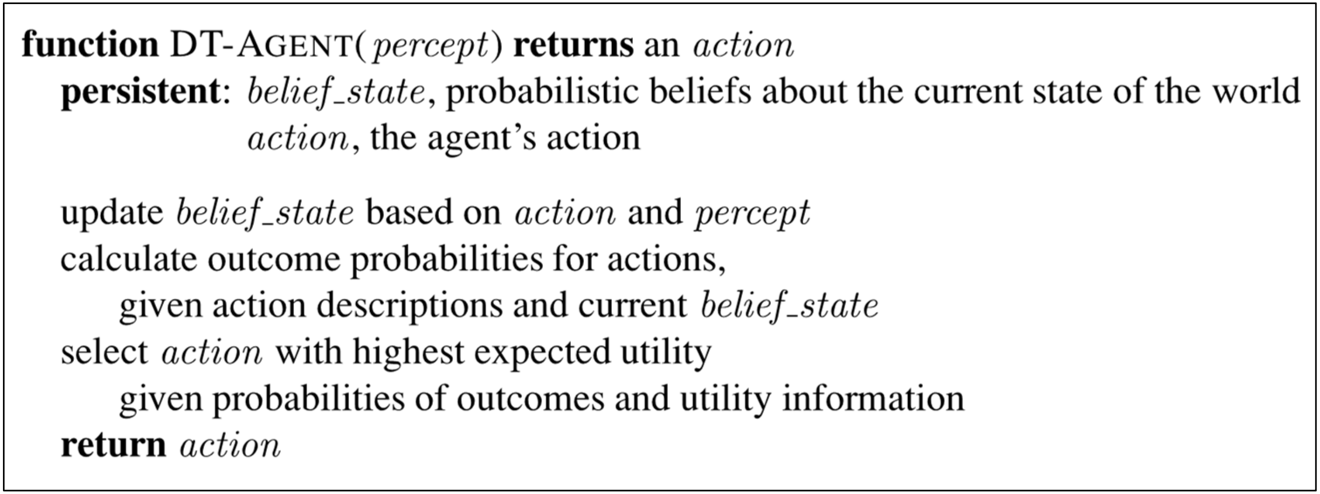
میانگین روی همه‌ی برآمدهای ممکن آن کنش

اصل حداکثر امید سودمندی
maximum expected utility (MEU)

عامل نظریه تصمیمی

DECISION THEORETIC AGENT (DT-AGENT)

یک عامل نظریه تصمیمی که کنش‌های رسیونال را انتخاب می‌کند.



```
function DT-AGENT(percept) returns an action
    persistent: belief_state, probabilistic beliefs about the current state of the world
                action, the agent's action

    update belief_state based on action and percept
    calculate outcome probabilities for actions,
        given action descriptions and current belief_state
    select action with highest expected utility
        given probabilities of outcomes and utility information
    return action
```

حالت باور عامل نظریه تصمیمی نه تنها **امکان‌ها**، بلکه **احتمال‌های** حالت‌های دنیا را بازنمایی می‌کند.

probabilities **possibilities**

کمی‌سازی عدم اطمینان

۲

مبانی نمادگذاری احتمال

پایه‌های احتمال

PROBABILITY BASICS

فضای نمونه *Sample Space*	Ω	مجموعه‌ی همه‌ی برآمدهای ممکن (مثلاً ۶ برآمد ممکن در انداختن یک تاس)

نقطه‌ی نمونه *Sample Point*	**دنیای ممکن** *Possible World*	**رویداد اتمیک** *Atomic Event*	$\omega \in \Omega$	یکی از برآمدها (مثلاً برآمد ۴ در انداختن یک تاس)

فضای احتمال *Probability Space*	**مدل احتمال** *Probability Model*	(Ω, P)	یک فضای نمونه + انتساب احتمال به هر نقطه‌ی نمونه

an assignment $P(\omega)$ for every $\omega \in \Omega$ s.t.

$$0 \le P(\omega) \le 1$$

$$\Sigma_\omega P(\omega) = 1$$

مثلاً : $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6.$

رویداد *Event*	**پیشامد** *Event*	A	هر زیرمجموعه‌ای از فضای نمونه Ω

$$P(A) = \Sigma_{\{\omega \in A\}} P(\omega)$$

مثلاً : $P(\text{die roll} < 4) = P(1) + P(2) + P(3) = 1/6 + 1/6 + 1/6 = 1/2$

متغیرهای تصادفی

RANDOM VARIABLES

متغیر تصادفی *Random Variable*	X	تابعی از فضای نمونه به یک برد (مثلاً اعداد حقیقی، یا بولی)

مثلاً : $Odd(1) = true.$

توزیع احتمال *Probability Distribution*	هر متغیر تصادفی X یک تابع توزیع احتمال دارد.

$$P(X = x_i) = \Sigma_{\{\omega:X(\omega)=x_i\}} P(\omega)$$

مثلاً : $P(Odd = true) = P(1) + P(3) + P(5) = 1/6 + 1/6 + 1/6 = 1/2$

گزاره‌ها به‌عنوان رویدادها

به هر گزاره‌ی منطقی به عنوان یک رویداد (مجموعه‌ای از نقاط نمونه) نگاه می‌کنیم
که در آنها گزاره **true** است.

Given Boolean random variables A and B:

event a = set of sample points where $A(\omega) = true$

event $\neg a$ = set of sample points where $A(\omega) = false$

event $a \wedge b$ = points where $A(\omega) = true$ and $B(\omega) = true$

در کاربردهای هوش مصنوعی،
نقاط نمونه به‌وسیله‌ی مقادیر مجموعه‌ای از متغیرهای تصادفی **تعریف** می‌شود؛
یعنی: فضای نمونه ضرب دکارتی بردهای متغیرهاست.

برای متغیرهای بولی، نقطه‌ی نمونه = مدل منطق گزاره‌ای

e.g., $A = true$, $B = false$, or $a \wedge \neg b$.

گزاره = فصل رویدادهای اتمیک که **true** هستند.

e.g., $(a \vee b) \equiv (\neg a \wedge b) \vee (a \wedge \neg b) \vee (a \wedge b)$

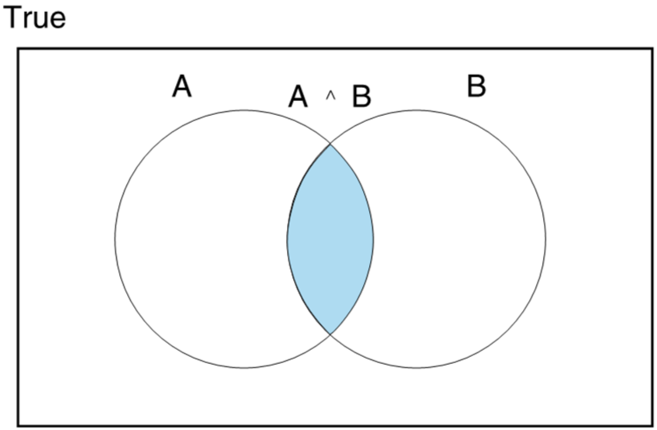
$$\Rightarrow \; P(a \vee b) = P(\neg a \wedge b) + P(a \wedge \neg b) + P(a \wedge b)$$

چرا از احتمال استفاده می‌کنیم؟

تعریف احتمال ایجاب می‌کند که رویدادهای خاص مرتبط به صورت منطقی، باید احتمال‌های مرتبط داشته باشند.

مثلاً:

$$P(a \vee b) = P(a) + P(b) - P(a \wedge b)$$



نحو برای گزاره‌ها

متغیر تصادفی گزاره‌ای *Propositional Random Variable*	**متغیر تصادفی بولی** *Boolean Random Variable*

e.g., $Cavity$ (do I have a cavity?)

$Cavity = true$ is a proposition, also written $cavity$

متغیر تصادفی گسسته *Discrete Random Variable*	متناهی **(finite)** نامتناهی **(infinite)**

e.g., $Weather$ is one of $\langle sunny, rain, cloudy, snow\rangle$

$Weather = rain$ is a proposition

مقادیر، باید جامع و دو به دو متمایز (مانعة الجمع) باشند.

متغیر تصادفی پیوسته *Continuous Random Variable*	کران‌دار **(bounded)** بی‌کران **(unbounded)**

e.g., $Temp = 21.6$; also allow, e.g., $Temp < 22.0$.

هر ترکیب بولی دلخواه از گزاره‌های پایه قابل استفاده است.

احتمال پیشین

PRIOR PROBABILITY

احتمال پیشین *Prior Probability*	احتمال غیرشرطی *Unconditional Probability*

e.g., $P(Cavity = true) = 0.1$ and $P(Weather = sunny) = 0.72$

احتمال پیشین هر گزاره، متناظر باور پیشین به آمدن هر شاهد (جدید) است.

توزیع احتمال توأم

JOINT PROBABILITY DISTRIBUTION

توزیع احتمال
Probability Distribution

مقادیر احتمال را برای همه‌ی انتساب‌های ممکن به متغیر تصادفی مشخص می‌کند.

$\mathbf{P}(Weather) = \langle 0.72, 0.1, 0.08, 0.1 \rangle$ (normalized, i.e., sums to 1)

توزیع احتمال توأم
Joint Probability Distribution

برای یک مجموعه متغیر تصادفی، احتمال هر رویداد اتمیک بر روی آن متغیرهای تصادفی را مشخص می‌کند.

$\mathbf{P}(Weather, Cavity)$ = a 4×2 matrix of values:

$Weather =$	$sunny$	$rain$	$cloudy$	$snow$
$Cavity = true$	0.144	0.02	0.016	0.02
$Cavity = false$	0.576	0.08	0.064	0.08

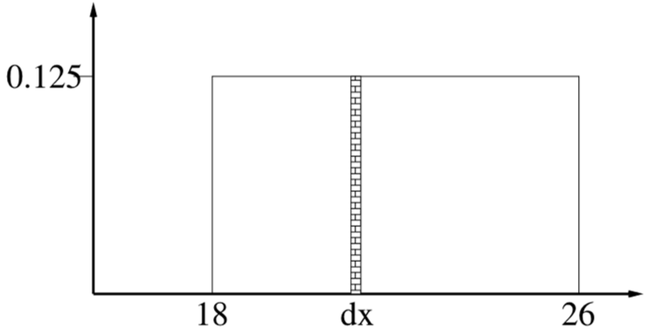
هر پرسشی در مورد یک دامنه، می‌تواند به کمک توزیع توأم آن پاسخ داده شود،
زیرا هر رویداد، مجموعی از نقاط نمونه است.

احتمال برای متغیرهای پیوسته

PROBABILITY FOR CONTINUOUS VARIABLES

برای متغیرهای پیوسته، توزیع احتمال به‌صورت یک تابع پارامتری از یک مقدار بیان می‌شود.

$P(X = x) = U[18, 26](x)$ = چگالی یکنواخت بین 18 و 26



در اینجا P، یک تابع چگالی احتمال (density) است که انتگرال آن ۱ می‌شود.

معنای

$$P(X = 20.5) = 0.125$$

در واقع عبارت است از:

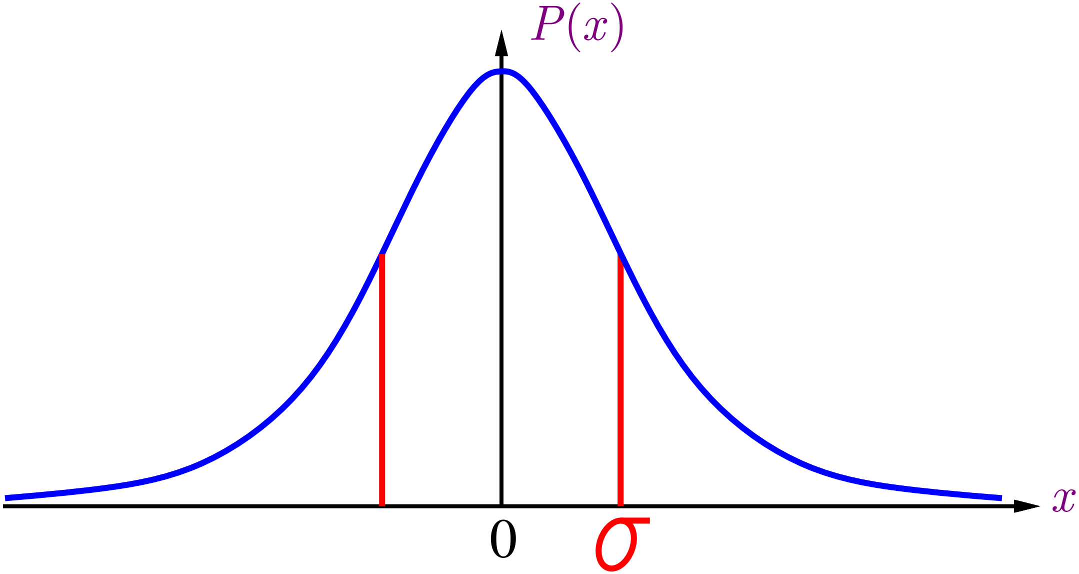
$$\lim_{dx \to 0} P(20.5 \leq X \leq 20.5 + dx) \,/\, dx = 0.125$$

احتمال برای متغیرهای پیوسته

مثال: تابع چگالی گاوسی

GAUSSIAN DENSITY

$$P(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)^2/2\sigma^2}$$



احتمال شرطی

CONDITIONAL PROBABILITY

احتمال پسین *Posterior Probability*	**احتمال شرطی** *Conditional Probability*

e.g., $P(cavity|toothache) = 0.8$

به معنی احتمال با فرض اینکه «$toothace$ **همه‌ی آن چیزی است که من می‌دانم**»
نه اینکه «اگر $toothace$ باشد، آن‌گاه 80% شانس $cavity$ وجود دارد»

نمادگذاری برای توزیع‌های شرطی:

$\mathbf{P}(Cavity|Toothache)$ = 2-element vector of 2-element vectors

اگر بیشتر بدانیم، مثلاً $cavity$ هم داده شده باشد، در این صورت داریم:

$$P(cavity|toothache, cavity) = 1$$

تذکر: باوری که کمتر خاص است، پس از رسیدن شاهد بیشتر، **معتبر باقی می‌ماند**؛
اما این همیشه **مفید** نیست.

شاهد جدید، ممکن است نامربوط باشد؛ در این صورت می‌توان ساده‌سازی کرد، مثلاً:

$$P(cavity|toothache, 49ersWin) = P(cavity|toothache) = 0.8$$

احتمال شرطی

تعریف

CONDITIONAL PROBABILITY

قاعده‌ی بیز
Bayes' Rule

$$P(a|b) = \frac{P(a \wedge b)}{P(b)} \text{ if } P(b) \neq 0$$

قاعده‌ی ضرب
Product Rule

فرمول‌بندی جایگزین برای قاعده‌ی بیز:

$$P(a \wedge b) = P(a|b)P(b) = P(b|a)P(a)$$

نسخه‌ی عمومی برای کل توزیع، به‌صورت زیر است:

$$\mathbf{P}(Weather, Cavity) = \mathbf{P}(Weather|Cavity)\mathbf{P}(Cavity)$$

(View as a 4×2 set of equations, **not** matrix mult.)

قاعده‌ی زنجیره‌ای
Chain Rule

استخراج با به‌کارگیری پی‌درپی قاعده‌ی ضرب:

$$\begin{aligned}\mathbf{P}(X_1, \ldots, X_n) &= \mathbf{P}(X_1, \ldots, X_{n-1})\ \mathbf{P}(X_n|X_1, \ldots, X_{n-1}) \\ &= \mathbf{P}(X_1, \ldots, X_{n-2})\ \mathbf{P}(X_{n-1}|X_1, \ldots, X_{n-2})\ \mathbf{P}(X_n|X_1, \ldots, X_{n-1}) \\ &= \ldots \\ &= \Pi_{i=1}^{n} \mathbf{P}(X_i|X_1, \ldots, X_{i-1})\end{aligned}$$

۳

استنتاج
با استفاده
از
توزیع توأم
کامل

استنتاج با شمارش

مثال (۱ از ۵)

INFERENCE BY ENUMERATION

شروع با توزیع توأم:

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ *cavity*	**.016**	**.064**	**.144**	**.576**

برای هر گزاره ϕ، احتمال رویدادهای اتمیک که ϕ در آنها درست است را جمع کنید:

$$P(\phi) = \Sigma_{\omega:\omega\models\phi} P(\omega)$$

استنتاج با شمارش

مثال (۲ از ۵)

INFERENCE BY ENUMERATION

شروع با توزیع توأم:

	toothache		$\neg$ *toothache*	
	catch	$\neg$ *catch*	*catch*	$\neg$ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
$\neg$ *cavity*	**.016**	**.064**	**.144**	**.576**

برای هر گزاره ϕ، احتمال رویدادهای اتمیک که ϕ در آنها درست است را جمع کنید:

$$P(\phi) = \Sigma_{\omega:\omega\models\phi} P(\omega)$$

$$P(toothache) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2$$

استنتاج با شمارش

مثال (۳ از ۵)

INFERENCE BY ENUMERATION

شروع با توزیع توأم:

	toothache		¬ toothache	
	catch	¬ catch	catch	¬ catch
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

برای هر گزاره ϕ، احتمال رویدادهای اتمیک که ϕ در آنها درست است را جمع کنید:

$$P(\phi) = \Sigma_{\omega:\omega\models\phi} P(\omega)$$

$$P(cavity \vee toothache) = 0.108+0.012+0.072+0.008+0.016+0.064 = 0.28$$

استنتاج با شمارش

مثال (۴ از ۵)

INFERENCE BY ENUMERATION

شروع با توزیع توأم:

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ *cavity*	**.016**	**.064**	**.144**	**.576**

می‌توانیم احتمال‌های شرطی را نیز محاسبه کنیم:

$$
\begin{aligned}
P(\neg cavity | toothache) &= \frac{P(\neg cavity \wedge toothache)}{P(toothache)} \\
&= \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4
\end{aligned}
$$

استنتاج با شمارش

مثال (۵ از ۵)

INFERENCE BY ENUMERATION

شروع با توزیع توأم:

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ *cavity*	**.016**	**.064**	**.144**	**.576**

مخرج کسر می‌تواند به‌عنوان یک ثابت نرمال‌سازی α دیده شود:

$$\begin{aligned}\mathbf{P}(Cavity|toothache) &= \alpha\,\mathbf{P}(Cavity, toothache)\\ &= \alpha\,[\mathbf{P}(Cavity, toothache, catch) + \mathbf{P}(Cavity, toothache, \neg catch)]\\ &= \alpha\,[\langle 0.108, 0.016\rangle + \langle 0.012, 0.064\rangle]\\ &= \alpha\,\langle 0.12, 0.08\rangle = \langle 0.6, 0.4\rangle\end{aligned}$$

ایده‌ی عمومی: توزیع را بر روی متغیر پرسش محاسبه کنید:
با ثابت کردن متغیرهای شاهد و محاسبه‌ی مجموع روی متغیرهای پنهان

استنتاج با شمارش

قاعده‌ی کلی

INFERENCE BY ENUMERATION

فرض کنیم $\mathbf{X}$ همه‌ی متغیرها باشد. معمولاً می‌خواهیم
توزیع توأم پسین را برای متغیرهای پرسش $\mathbf{Y}$ به دست آوریم
با داشتن مقادیر خاص $\mathbf{e}$ برای متغیرهای شاهد $\mathbf{E}$

متغیر پرس‌وجو
Query Variable

متغیر شاهد
Evidence Variable

متغیر پنهان
Hidden Variable

متغیرهای پنهان می‌شود:

$$\mathbf{H} = \mathbf{X} - \mathbf{Y} - \mathbf{E}$$

در این صورت:
مجموع‌یابی لازم برای درایه‌های توأم با مجموع‌گیری روی متغیرهای پنهان انجام می‌شود:

$$\mathbf{P}(\mathbf{Y}|\mathbf{E}=\mathbf{e}) = \alpha \mathbf{P}(\mathbf{Y},\mathbf{E}=\mathbf{e}) = \alpha \Sigma_{\mathbf{h}} \mathbf{P}(\mathbf{Y},\mathbf{E}=\mathbf{e},\mathbf{H}=\mathbf{h})$$

جملات موجود در مجموع درایه‌های توأم هستند،
زیرا $\mathbf{Y}$، $\mathbf{E}$ و $\mathbf{H}$ با هم مجموعه‌ی کل متغیرهای تصادفی را پوشش می‌دهند.

استنتاج با شمارش

مشکلات

INFERENCE BY ENUMERATION

۱) پیچیدگی زمانی بدترین حالت $O(d^n)$

۲) پیچیدگی فضایی برای ذخیره‌سازی توزیع توأم $O(d^n)$

۳) چگونه باید اعداد را برای $O(d^n)$ درایه پیدا کرد؟؟؟

d = بزرگترین اندازه‌ی چندتایی‌ها

Prepared by Kazim Fouladi | Spring 2017 | 3rd Edition

هوش مصنوعی

کمی‌سازی عدم اطمینان

۴

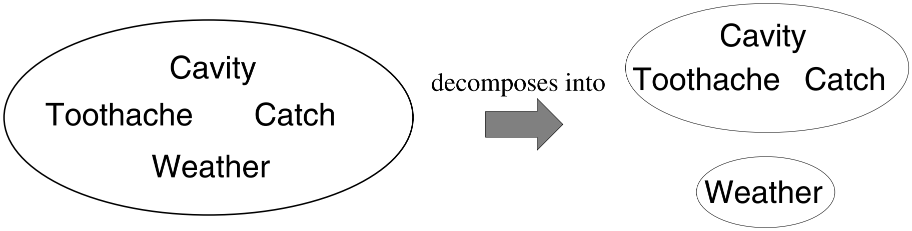
استقلال

استقلال

INDEPENDENCE

دو متغیر تصادفی A و B مستقل هستند، اگر و فقط اگر

$$\mathbf{P}(A|B)=\mathbf{P}(A) \quad \text{or} \quad \mathbf{P}(B|A)=\mathbf{P}(B) \quad \text{or} \quad \mathbf{P}(A,B)=\mathbf{P}(A)\mathbf{P}(B)$$



$$\mathbf{P}(Toothache, Catch, Cavity, Weather)$$
$$= \mathbf{P}(Toothache, Catch, Cavity)\mathbf{P}(Weather)$$

32 entries reduced to 12; for n independent biased coins, $2^n \rightarrow n$

استقلال مطلق قدرت‌مند است، اما به‌ندرت پیش می‌آید

دندان‌پزشکی یک حوزه‌ی بزرگ با صدها متغیر است،
که هیچ یک از دیگری مستقل نیست ...
پس چه باید کرد؟ استفاده از مفهوم استقلال شرطی

Prepared by Kazim Fouladi | Spring 2017 | 3rd Edition

کاظم فولادی

استقلال شرطی

مثال

CONDITIONAL INDEPENDENCE

$\mathbf{P}(Toothache, Cavity, Catch)$ has $2^3 - 1 = 7$ independent entries

اگر بیمار کرم‌خوردگی $(cavity)$ داشته باشد،
احتمال اینکه میله در آن گیر $(catch)$ کند، به اینکه بیمار دندان‌درد $(toothace)$ داشته باشد یا نه، بستگی ندارد:

$$(1)\ P(catch|toothache, cavity) = P(catch|cavity)$$

شبیه همین استقلال برقرار است، اگر بیمار کرم‌خوردگی $(cavity)$ نداشته باشد:

$$(2)\ P(catch|toothache, \neg cavity) = P(catch|\neg cavity)$$

$Catch$ is conditionally independent of $Toothache$ given $Cavity$:

$$\mathbf{P}(Catch|Toothache, Cavity) = \mathbf{P}(Catch|Cavity)$$

استقلال شرطی یک متغیر از یک متغیر دیگر، با داشتن (به شرط) یک متغیر دیگر

جملات معادل:

$$\mathbf{P}(Toothache|Catch, Cavity) = \mathbf{P}(Toothache|Cavity)$$

$$\mathbf{P}(Toothache, Catch|Cavity) = \mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)$$

استقلال شرطی

مثال

CONDITIONAL INDEPENDENCE

نوشتن توزیع توأم کامل با استفاده از قاعده‌ی زنجیره‌ای:

$$
\begin{aligned}
&\mathbf{P}(Toothache, Catch, Cavity) \\
&= \mathbf{P}(Toothache|Catch, Cavity)\mathbf{P}(Catch, Cavity) \\
&= \mathbf{P}(Toothache|Catch, Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity) \\
&= \mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity)
\end{aligned}
$$

I.e., $2 + 2 + 1 = 5$ independent numbers (equations 1 and 2 remove 2)

در بیشتر موارد، استفاده از استقلال شرطی، اندازه‌ی بازنمایی توزیع توأم n متغیر را از اندازه‌ی نمایی بر حسب n به اندازه‌ی خطی بر حسب n کاهش می‌دهد.

استقلال شرطی، پایه‌ای‌ترین و مقاوم‌ترین صورت از دانایی ما در مورد محیط‌های نامطمئن است.

۵

قاعده‌ی
بیز
و
کاربرد آن

قاعده‌ی بیز

BAYES' RULE

قاعده‌ی بیز
Bayes' Rule

Product rule $P(a \land b) = P(a|b)P(b) = P(b|a)P(a)$

$\Rightarrow$ Bayes' rule $P(a|b) = \dfrac{P(b|a)P(a)}{P(b)}$

در فرم توزیع:

$$\mathbf{P}(Y|X) = \frac{\mathbf{P}(X|Y)\mathbf{P}(Y)}{\mathbf{P}(X)} = \alpha\mathbf{P}(X|Y)\mathbf{P}(Y)$$

مفید برای سنجش احتمال تشخیصی از روی احتمال علّی

$$P(Cause|Effect) = \frac{P(Effect|Cause)P(Cause)}{P(Effect)}$$

برای مثال: اگر M بیماری منتزیت و S گردن‌درد باشد:

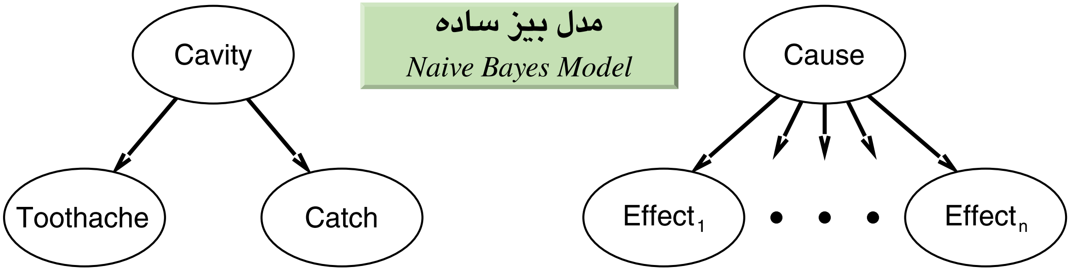
$$P(m|s) = \frac{P(s|m)P(m)}{P(s)} = \frac{0.8 \times 0.0001}{0.1} = 0.0008$$

قاعده‌ی بیز و استقلال شرطی

$$\begin{aligned}&\mathbf{P}(Cavity|toothache \wedge catch)\\ &= \alpha\,\mathbf{P}(toothache \wedge catch|Cavity)\mathbf{P}(Cavity)\\ &= \alpha\,\mathbf{P}(toothache|Cavity)\mathbf{P}(catch|Cavity)\mathbf{P}(Cavity)\end{aligned}$$

این مثالی است از مدل بیز ساده

$$\mathbf{P}(Cause, Effect_1, \ldots, Effect_n) = \mathbf{P}(Cause)\Pi_i\mathbf{P}(Effect_i|Cause)$$



مدل بیز ساده

Naive Bayes Model

تعداد کل پارامترهای لازم، بر حسب تعداد متغیرها (n) **خطی** است.

Prepared by Kazim Fouladi | Spring 2017 | 3rd Edition

هوش مصنوعی

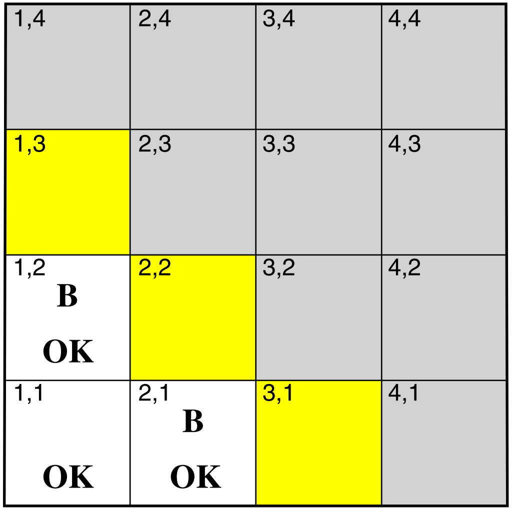
کمی‌سازی عدم اطمینان

۶

بازبینی دنیای اژدها

دنیای اژدها

WUMPUS WORLD



1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 **B** **OK**	2,2	3,2	4,2
1,1 **OK**	2,1 **B** **OK**	3,1	4,1

$P_{ij} = true$ iff $[i, j]$ contains a pit

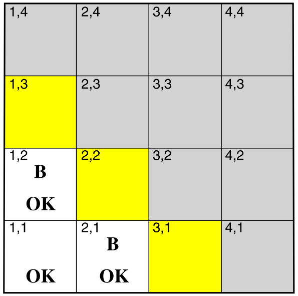
$B_{ij} = true$ iff $[i, j]$ is breezy

Include only $B_{1,1}, B_{1,2}, B_{2,1}$ in the probability model

دنیای اژدها

مشخص‌سازی مدل احتمال

WUMPUS WORLD



1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 B OK	2,2	3,2	4,2
1,1 OK	2,1 B OK	3,1	4,1

$P_{ij} = true$ iff $[i,j]$ contains a pit

$B_{ij} = true$ iff $[i,j]$ is breezy

Include only $B_{1,1}, B_{1,2}, B_{2,1}$ in the probability model

The full joint distribution is $\mathbf{P}(P_{1,1}, \ldots, P_{4,4}, B_{1,1}, B_{1,2}, B_{2,1})$

Apply product rule: $\mathbf{P}(B_{1,1}, B_{1,2}, B_{2,1} \mid P_{1,1}, \ldots, P_{4,4})\mathbf{P}(P_{1,1}, \ldots, P_{4,4})$

(Do it this way to get $P(Effect|Cause)$.)

☞ First term: 1 if pits are adjacent to breezes, 0 otherwise

☞ Second term: pits are placed randomly, probability 0.2 per square:

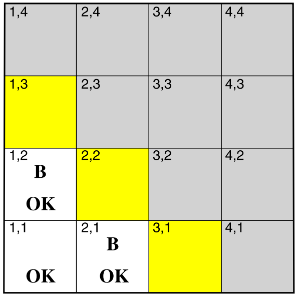
$$\mathbf{P}(P_{1,1}, \ldots, P_{4,4}) = \Pi_{i,j=1,1}^{4,4} \mathbf{P}(P_{i,j}) = 0.2^n \times 0.8^{16-n}$$

for n pits.

دنیای اژدها

مشاهدات و پرسش‌ها

WUMPUS WORLD



1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 B OK	2,2	3,2	4,2
1,1 OK	2,1 B OK	3,1	4,1

We know the following facts:

$$b = \neg b_{1,1} \wedge b_{1,2} \wedge b_{2,1}$$
$$known = \neg p_{1,1} \wedge \neg p_{1,2} \wedge \neg p_{2,1}$$

Query is $\mathbf{P}(P_{1,3}|known, b)$

Define $Unknown = P_{ij}$s other than $P_{1,3}$ and $Known$

For inference by enumeration, we have

$$\mathbf{P}(P_{1,3}|known, b) = \alpha \Sigma_{unknown} \mathbf{P}(P_{1,3}, unknown, known, b)$$

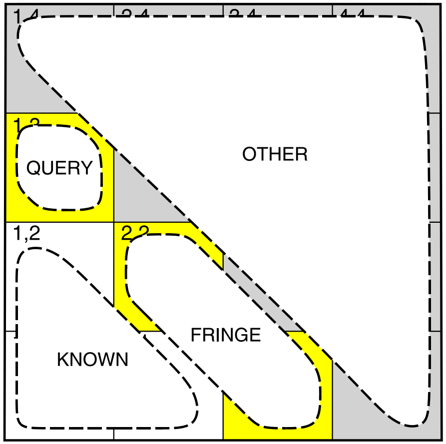
Grows exponentially with number of squares!

دنیای اژدها

استفاده از استقلال شرطی

WUMPUS WORLD

دید پایه: مشاهدات از سایر خانه‌های پنهان مستقل شرطی هستند
با داشتن خانه‌های همسایه‌ی پنهان



Define $Unknown = Fringe \cup Other$

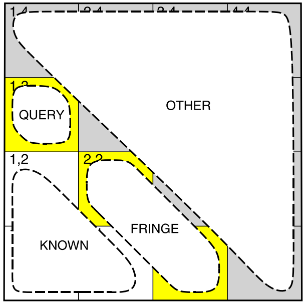
$$\mathbf{P}(b|P_{1,3}, Known, Unknown) = \mathbf{P}(b|P_{1,3}, Known, Fringe)$$

باید پرسش را به‌گونه‌ای دستکاری کنیم که بتوان از این فرمول استفاده کرد.

دنیای اژدها

استفاده از استقلال شرطی

WUMPUS WORLD



Define $Unknown = Fringe \cup Other$

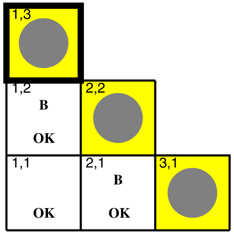
$\mathbf{P}(b|P_{1,3}, Known, Unknown) = \mathbf{P}(b|P_{1,3}, Known, Fringe)$

$$\begin{aligned}
\mathbf{P}(P_{1,3}|known, b) &= \alpha \sum_{unknown} \mathbf{P}(P_{1,3}, unknown, known, b) \\
&= \alpha \sum_{unknown} \mathbf{P}(b|P_{1,3}, known, unknown)\mathbf{P}(P_{1,3}, known, unknown) \\
&= \alpha \sum_{fringe} \sum_{other} \mathbf{P}(b|known, P_{1,3}, fringe, other)\mathbf{P}(P_{1,3}, known, fringe, other) \\
&= \alpha \sum_{fringe} \sum_{other} \mathbf{P}(b|known, P_{1,3}, fringe)\mathbf{P}(P_{1,3}, known, fringe, other) \\
&= \alpha \sum_{fringe} \mathbf{P}(b|known, P_{1,3}, fringe) \sum_{other} \mathbf{P}(P_{1,3}, known, fringe, other) \\
&= \alpha \sum_{fringe} \mathbf{P}(b|known, P_{1,3}, fringe) \sum_{other} \mathbf{P}(P_{1,3})P(known)P(fringe)P(other) \\
&= \alpha\, P(known)\mathbf{P}(P_{1,3}) \sum_{fringe} \mathbf{P}(b|known, P_{1,3}, fringe)P(fringe) \sum_{other} P(other) \\
&= \alpha'\, \mathbf{P}(P_{1,3}) \sum_{fringe} \mathbf{P}(b|known, P_{1,3}, fringe)P(fringe)
\end{aligned}$$

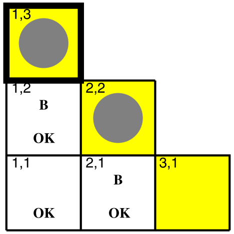
دنیای اژدها

استفاده از استقلال شرطی

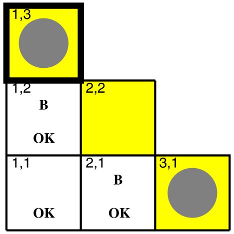
WUMPUS WORLD



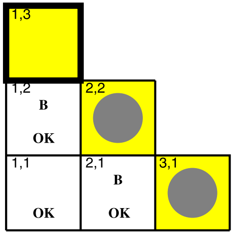
0.2 x 0.2 = 0.04



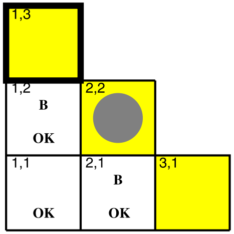
0.2 x 0.8 = 0.16



0.8 x 0.2 = 0.16



0.2 x 0.2 = 0.04



0.2 x 0.8 = 0.16

$$
\begin{aligned}
\mathbf{P}(P_{1,3}|known, b) &= \alpha' \langle 0.2(0.04 + 0.16 + 0.16),\ 0.8(0.04 + 0.16) \rangle \\
&\approx \langle 0.31, 0.69 \rangle
\end{aligned}
$$

$$
\mathbf{P}(P_{2,2}|known, b) \approx \langle 0.86, 0.14 \rangle
$$

Prepared by Kazim Fouladi | Spring 2017 | 3rd Edition

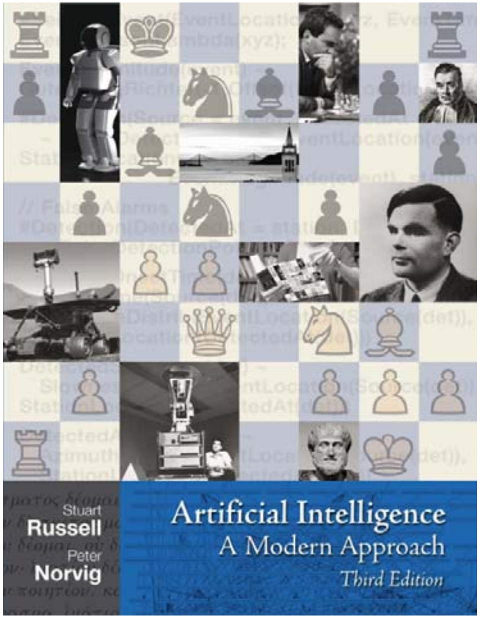
هوش مصنوعی

کمی‌سازی عدم اطمینان

۷

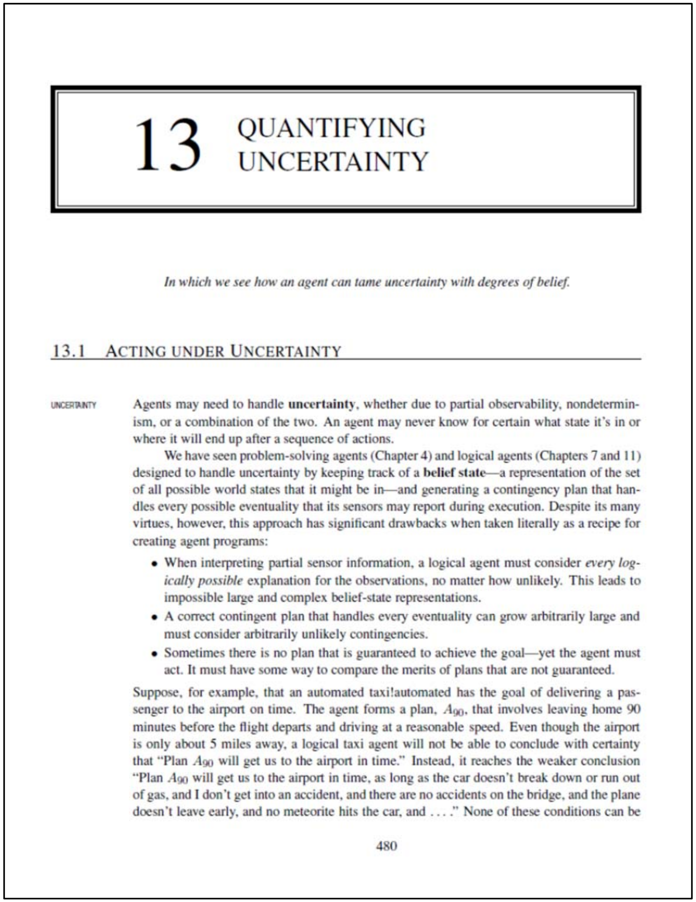
منابع، مطالعه، تکلیف

منبع اصلی



Stuart Russell and Peter Norvig,
Artificial Intelligence: A Modern Approach,
3rd Edition, Prentice Hall, 2010.

Chapter 13



13 QUANTIFYING UNCERTAINTY

In which we see how an agent can tame uncertainty with degrees of belief.

13.1 ACTING UNDER UNCERTAINTY

UNCERTAINTY

Agents may need to handle **uncertainty**, whether due to partial observability, nondeterminism, or a combination of the two. An agent may never know for certain what state it's in or where it will end up after a sequence of actions.

We have seen problem-solving agents (Chapter 4) and logical agents (Chapters 7 and 11) designed to handle uncertainty by keeping track of a **belief state**—a representation of the set of all possible world states that it might be in—and generating a contingency plan that handles every possible eventuality that its sensors may report during execution. Despite its many virtues, however, this approach has significant drawbacks when taken literally as a recipe for creating agent programs:

- When interpreting partial sensor information, a logical agent must consider *every logically possible* explanation for the observations, no matter how unlikely. This leads to impossible large and complex belief-state representations.
- A correct contingent plan that handles every eventuality can grow arbitrarily large and must consider arbitrarily unlikely contingencies.
- Sometimes there is no plan that is guaranteed to achieve the goal—yet the agent must act. It must have some way to compare the merits of plans that are not guaranteed.

Suppose, for example, that an automated taxi!automated has the goal of delivering a passenger to the airport on time. The agent forms a plan, A_{90}, that involves leaving home 90 minutes before the flight departs and driving at a reasonable speed. Even though the airport is only about 5 miles away, a logical taxi agent will not be able to conclude with certainty that "Plan A_{90} will get us to the airport in time." Instead, it reaches the weaker conclusion "Plan A_{90} will get us to the airport in time, as long as the car doesn't break down or run out of gas, and I don't get into an accident, and there are no accidents on the bridge, and the plane doesn't leave early, and no meteorite hits the car, and . . . ." None of these conditions can be

480