



تمرین شماره ۹

فصل دوازدهم و سیزدهم

کمی سازی عدم اطمینان و استدلال احتمالاتی

QUANTIFYING UNCERTAINTY AND PROBABILISTIC REASONING

۱) با توجه به توزیع توأم کامل داده شده، مقادیر زیر را محاسبه کنید.

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

(الف) $P(\text{toothache})$

(ب) $P(\text{cavity})$

(ج) $P(\text{Cavity})$

(د) $P(\text{Toothache}|\text{cavity})$

(ه) $P(\text{Cavity}|\text{toothache} \vee \text{catch})$

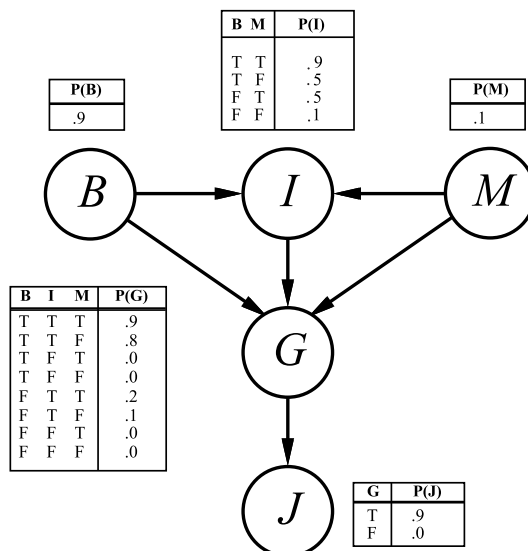
۲) پس از انجام یک چک‌آپ، پزشک یک خبر خوب و یک خبر بد برای شما دارد:

- خبر بد: تست شما مثبت بوده است؛ به این معنی که شما دچار یک بیماری جدی شده‌اید. از طرفی این تست ۹۹٪ دقیق است (احتمال اینکه شما بیمار باشید و جواب تست مثبت باشد = احتمال اینکه شما بیمار نباشید و جواب تست منفی باشد = ۹۹٪).

- خبر خوب: این بیماری یک بیماری نادر است که از بین ۱۰۰۰۰ نفر، فقط ۱ نفر به آن مبتلا می‌شود.

چرا این یک خبر خوب است؟ شانس اینکه شما واقعاً این بیماری را داشته باشید، چه قدر است؟

۳) شبکه‌ی بیزی زیر را در نظر بگیرید:



(الف) بدون در نظر گرفتن CPTها، کدامیک از عبارتهای زیر از ساختار شبکه نتیجه می شود؟

(i) $P(B, I, M) = P(B)P(I)P(M)$

(ii) $P(J|G) = P(J|G, I)$

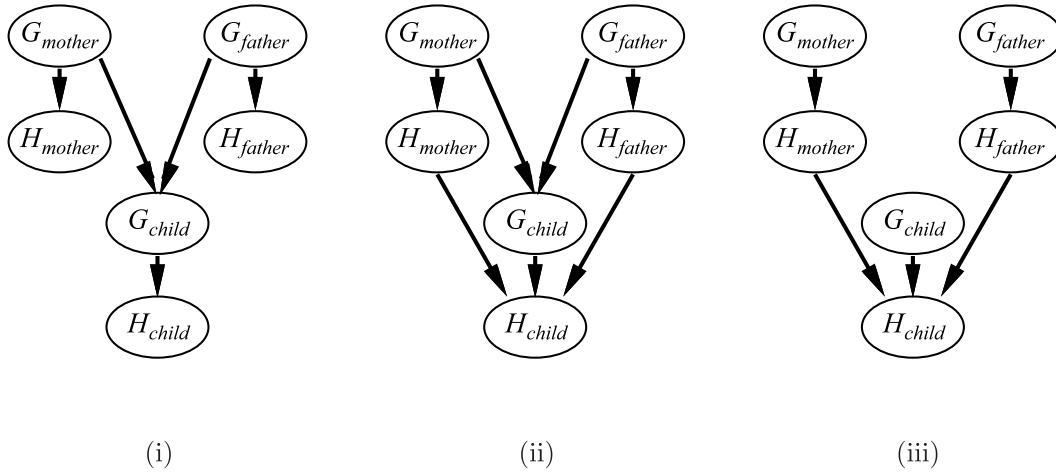
(iii) $P(M|G, B, I) = P(M|G, B, I, J)$

(ب) مقدار $P(b, i, -m, g, j)$ را محاسبه کنید.

(۴) فرض کنید H_x یک متغیر تصادفی باشد که به راستدستی یا چپدستی (Handness) شخص x با مقادیر ممکن r یا l اشاره دارد. یک فرض معمول این است که راستدستی یا چپدستی با یک مکانیزم ساده به ارث برده می شود، یعنی احتمالاً ژن G_x وجود دارد که با مقادیر ممکن r یا l و با احتمال s موجب راستدستی یا چپدستی می شود. به علاوه، خود این ژن احتمالاً باید از یکی از والدین شخص با احتمال کوچک m برای جهش در راستدستی یا چپدستی به ارث برده شده باشد.

(الف) کدامیک از شبکه های بیزی زیر ادعا را بیان می کند؟

$$P(G_{father}, G_{mother}, G_{child}) = P(G_{father})P(G_{mother})P(G_{child})$$



(ب) کدامیک از این شبکه ها ادعاهای استقلال سازگار با فرضیهی صورت مسئله را در بردارد؟

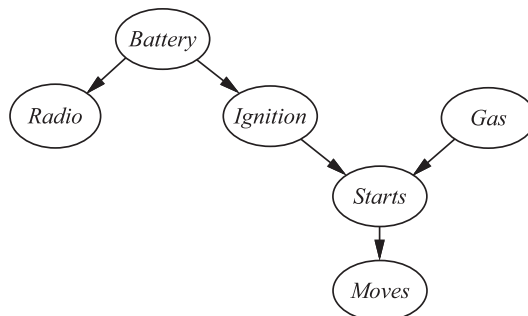
(ج) کدامیک از این شبکه ها بهترین توصیف فرضیهی صورت مسئله است؟

(د) CPT را برای $P(G_{child})$ در شبکه های (i) یا (ii) رسم کنید و مقادیر آن را پر کنید (جهش های $l \rightarrow r$ و $r \rightarrow l$ وقتی والد ها ژن های مختلف دارند، لغو می شود).

(ه) فرض کنید که $P(G_{father} = l) = P(G_{mother} = l) = x$. در شبکه های (i) یا (ii) عبارتی را برای $P(G_{child} = l)$ بر حسب x و m تنها، با شرط گذاری روی گره های والد آن استخراج کنید.

(و) تحت شرایط تعادل ژنتیکی انتظار داریم که توزیع ژن ها در خلال نسل ها مشابه باشد. با استفاده از این نکته، مقدار x را محاسبه کنید و با علم به وضعیت راستدستی-چپدستی در انسان ها توضیح دهید که چرا فرضیهی توصیف شده در ابتدای این مسئله بایستی نادرست باشد؟

(۵) شبکه ی بیزی زیر را برای عیب یابی خودرو در نظر بگیرید:



- (الف) این شبکه را با متغیرهای بولی *IcyWeather* (هوای یخ) و *StarterMotor* (موتور استارت) گسترش دهید.
- (ب) برای همه‌ی گره‌ها، جدول‌های احتمال شرطی با مقادیر نوعی مستدل ارائه دهید.
- (ج) با فرض اینکه هیچ رابطه‌ی استقلال شرطی بین متغیرهای تصادفی وجود نداشته باشد، چند مقدار مستقل در توزیع توأم احتمال ۸ متغیر بولی وجود دارد؟
- (د) چند مقدار احتمال مستقل در جدول‌های شبکه‌ی بیزی فوق وجود دارد؟
- (ه) [اختیاری] توزیع شرطی برای *Starts* می‌تواند به صورت یک توزیع noisy-AND توصیف شود. این خانواده مدل را در حالت کلی تعریف کنید و آن را به توزیع noisy-OR مرتبط کنید.
- ۶ [اختیاری] در یک سایت انرژی هسته‌ای، زنگ هشداری وجود دارد که در هنگام تجاوز درجه‌ی دما از یک مقدار آستانه‌ی داده شده به صدا درمی‌آید. درجه، دمای هسته را اندازه‌گیری می‌کند. متغیرهای بولی A (صدای زنگ خطر)، F_A (خراب بودن زنگ خطر)، F_G (خراب بودن درجه)، متغیرهای چندمقداری G (مقدار خوانده شده از درجه) و T (دمای واقعی هسته) را در نظر بگیرید.
- (الف) یک شبکه‌ی بیزی برای این دامنه رسم کنید، با توجه به اینکه درجه به احتمال بیشتر خراب می‌شود، هرگاه دمای هسته بسیار بالا رود.
- (ب) آیا این شبکه چنددرختی (polytree) (شبکه‌ای که حداکثر یک مسیر بدون جهت بین هر دو گره‌ی آن وجود دارد) است؟
- (ج) فرض کنید که تنها دو حالت دمای اندازه‌گیری شده‌ی واقعی وجود دارد: عادی (normal) و بالا (high). احتمال اینکه درجه، دمای درستی را نشان دهد هنگامی که کار می‌کند، x و احتمال اینکه درجه، دمای درستی را نشان دهد هنگامی که خراب است، y است. جدول احتمال شرطی متناظر با G را رسم کنید.
- (د) فرض کنید که زنگ خطر به درستی کار می‌کند، مگر اینکه خراب باشد. در این حالت صدایی نخواهد داشت. جدول احتمال شرطی متناظر با A را رسم کنید.
- (ه) فرض کنید که زنگ خطر و درجه هر دو کار می‌کنند و زنگ خطر به صدا درمی‌آید. عبارتی را برای «احتمال اینکه دمای هسته بسیار بالا است» بر حسب احتمالات شرطی مختلف در این شبکه محاسبه کنید.
- ۷ [اختیاری] دسته‌بندی متون (Text Categorization) عمل انتساب یک سند متنی داده شده به یکی از دسته‌های مشخص بر اساس محتوای متن است. برای این کار اغلب از مدل (Naive Bayes) استفاده می‌شود. در این مدل‌ها، متغیر پرس‌وجو (query variable) دسته‌ی سند و متغیرهای اثر (effect) حضور یا غیاب هر یک از کلمات زبان است. فرض این است که کلمات به‌طور مستقل در اسناد ظاهر می‌شوند که فراوانی آنها توسط دسته‌ی سند تعیین می‌شود.
- (الف) به‌طور دقیق تشریح کنید که چگونه چنین مدلی می‌تواند ساخته شود؟ فرض کنید که داده‌های آموزشی (training data) به‌صورت مجموعه‌ای از اسناد که تاکنون به دسته‌ها نسبت‌دهی شده‌اند، داده شده باشد.
- (ب) به‌طور دقیق توضیح دهید که چگونه یک سند جدید دسته‌بندی می‌شود؟
- (ج) آیا فرض استقلال فرض مستدل و درستی است؟ توضیح دهید؟