به نام خداوند بخشنده مهربان

هوش مصنوعی

دانشکده مهندسی کامپیوتر

بهار۱۴۰۴

۱۰ اردیبهشت ۱۴۰۱ | **میانترم درس هوش مصنوعی** | زمان آزمون: ۱۵۰ دقیقه

۱. لطفا پاسخ خود را با خط خوانا بنویسید.

۲. پاسخ هر سوال را در یک صفحه جدا و شماره پرسش را به صورت واضح در بالای هر صفحه بنویسید.

۳. آزمون از ۱۱۰ نمره است. دریافت ۱۰۰ نمره از ۱۱۰ نمره به منزله دریافت نمره کامل خواهد بود. نمره بالای ۱۰۰ سرریز نخواهد کرد.

۴. نوشته‌های شما در قسمت چرک‌نویس به هیچ عنوان تصحیح نخواهد شد.

۵. استفاده از منابع و لوازم الکترونیکی حین پاسخگویی به سوالات آزمون غیرمجاز است.

پرسش‌های آزمون (۱۱۰ نمره)

پرسش ۱ (۲۱ نمره) در هر مورد، درستی یا نادرستی عبارت را با توضیح مشخص کنید.

(آ) الگوریتم local beam search با $k = ۱$ معادل الگوریتم hill climbing عمل می کند.

(ب) در الگوریتم hill climbing with random restart با فرض متوسط ۳ حرکت تا رسیدن به شکست و ۴ حرکت تا به رسیدن موفقیت در هر مرحله و احتمال موفقیت $p = ۰/۱$ در هر مرحله ، امید ریاضی تعداد حرکات انجام شده برابر است با ۳۱

(ج) اگر گراف محدودیت ها در CSP یک دور داشته باشد ،مسئله در زمان چند جمله ای حل می شود.

(د) فرض کنید میخواهیم تابع fitness رشته ای را بیشینه کنیم و برای این کار از الگوریتم ژنتیک استفاده کنیم. و فرض کنید که عملیت های mutation و crossover داریم. در این صورت حتما به global optima می رسیم.

(ه) اگر h(s) admissible باشد، آنگاه $\sqrt{h(s)}$ و $max(۰, \log_۲ h(s))$ هم admissible هستند.

(و) الگوریتم uniform cost search با وزن یال های $-\alpha$ که $\alpha > ۰$ معادل الگوریتم DFS است.

(ز) اگر $g(s)$ و $f(s)$ دو تابع consistent باشند . میانگین آنها نیز consistent است.

پاسخ

(آ) **درست.** الگوریتم local beam search با k حالت کار می‌کند. در هر مرحله، جانشینان تمام k حالت را تولید کرده و k تا از بهترین جانشینان را برای مرحله بعد انتخاب می‌کند. اگر $k = ۱$ باشد، این الگوریتم تنها یک حالت را نگه می‌دارد، جانشینان آن را تولید می‌کند و بهترینِ آن‌ها را به عنوان حالت بعدی انتخاب می‌کند که این دقیقاً همان عملکرد الگوریتم hill climbing (از نوع steepest-ascent) است.

(ب) **درست.** این مسئله مربوط به توزیع هندسی است. امید ریاضی تعداد مراحل لازم برای رسیدن به اولین موفقیت برابر با $۱/p$ است. در اینجا $p = ۰/۱$ است، پس به طور متوسط به $۱/۰/۱ = ۱۰$ مرحله نیاز داریم. از این ۱۰ مرحله، به طور متوسط ۱ مرحله موفقیت‌آمیز و ۹ مرحله شکست خواهد بود. بنابراین امید ریاضی تعداد حرکات برابر است با:

$$E(\text{moves}) = (۹ \times \text{cost}_{\text{failure}}) + (۱ \times \text{cost}_{\text{success}}) = (۹ \times ۳) + (۱ \times ۴) = ۲۷ + ۴ = ۳۱$$

(ج) **درست.** اگر گراف محدودیت در یک مسئله CSP یک درخت باشد (یعنی هیچ دوری نداشته باشد)، مسئله در زمان چندجمله‌ای قابل حل است. اگر گراف فقط یک دور داشته باشد، می‌توانیم با مقداردهی به یک متغیر در آن دور، دور را بشکنیم. پس از مقداردهی به آن متغیر (امتحان کردن تمام مقادیر ممکن در دامنه آن)، گراف باقیمانده به یک درخت تبدیل می‌شود که به سرعت قابل حل است. بنابراین، کل مسئله در زمان چندجمله‌ای حل می‌شود.

(د) **نادرست.** الگوریتم ژنتیک یک الگوریتم جستجوی هیوریستیک و تصادفی است. اگرچه این الگوریتم اغلب در یافتن جواب‌های نزدیک به بهینه بسیار خوب عمل می‌کند، اما هیچ تضمینی برای رسیدن به بهینه سراسری (global optima) وجود ندارد. این الگوریتم ممکن است در یک بهینه محلی (local optima) همگرا شود.

(ه) **نادرست.** یک تابع هیوریستیک $h(s)$ قابل قبول (admissible) است اگر همیشه کمتر یا مساوی هزینه واقعی تا هدف ($h^*(s)$) باشد، یعنی $۰ \leq h(s) \leq h^*(s)$.

- برای $\sqrt{h(s)}$: این تابع لزوماً admissible نیست. به عنوان مثال، اگر $h^*(s) = ۰/۵$ و $h(s) = ۰/۳۶$ (که admissible است)، آنگاه $\sqrt{h(s)} = \sqrt{۰/۳۶} = ۰/۶$. در اینجا $\sqrt{h(s)} > h^*(s)$، بنابراین تابع جدید admissible نیست.

- برای $max(۰, \log_۲ h(s))$: این تابع admissible است. چون برای هر $x > ۰$ داریم $\log_۲(x) \leq x$. در نتیجه $\log_۲(h(s)) \leq h(s)$. از آنجایی که $h(s) \leq h^*(s)$، می‌توان نتیجه گرفت $\log_۲(h(s)) \leq h^*(s)$.

چون گزاره ادعا می‌کند که **هر دو** تابع admissible هستند، کل گزاره نادرست است.

(و) **درست.** الگوریتم uniform cost search همیشه گره‌ای را گسترش می‌دهد که کمترین هزینه مسیر ($g(n)$) را داشته باشد. اگر هزینه تمام یال‌ها یک عدد منفی ثابت مانند $-\alpha$ باشد، هزینه مسیر تا گره n در عمق d برابر با $-d \times \alpha$ خواهد بود. برای کمینه کردن این مقدار، الگوریتم باید گره‌ای با بیشترین عمق (d) را انتخاب کند. این دقیقاً همان استراتژی الگوریتم جستجوی اول عمق (DFS) است.

(ز) **درست.** یک تابع هیوریستیک h سازگار (consistent) است اگر برای هر گره s و جانشین آن s' داشته باشیم: $h(s) \leq \text{cost}(s, s') + h(s')$. فرض کنید $h_۱(s)$ و $h_۲(s)$ دو تابع سازگار باشند (با فرض اینکه منظور سوال از $g(s)$ و $f(s)$ دو تابع هیوریستیک مجزا بوده است). داریم:

$$h_۱(s) \leq \text{cost}(s, s') + h_۱(s')$$

$$h_۲(s) \leq \text{cost}(s, s') + h_۲(s')$$

با جمع کردن این دو نامساوی داریم:

$$h_۱(s) + h_۲(s) \leq ۲ \times \text{cost}(s, s') + h_۱(s') + h_۲(s')$$

با تقسیم طرفین بر ۲، به سادگی می‌توان دید که میانگین آن‌ها نیز در شرط سازگاری صدق می‌کند:

$$\frac{h_۱(s) + h_۲(s)}{۲} \leq \text{cost}(s, s') + \frac{h_۱(s') + h_۲(s')}{۲}$$

پرسش ۲ (۲۱ نمره)

شما در حال برنامه‌نویسی یک ربات خرید تعطیلات هستید که برای خرید تمام هدایای موجود در لیست خرید شما از یک فروشگاه به فروشگاه دیگر می‌رود. شما مجموعه‌ای از N هدیه $G = \{g_۱, g_۲, \ldots, g_N\}$ دارید که باید خریداری شوند. M فروشگاه $S = \{s_۱, s_۲, \ldots, s_M\}$ وجود دارد که هر کدام موجودی مشخصی از اقلام را در خود دارند: می گوییم $g_k \in s_i$ اگر فروشگاه s_i هدیه g_k را داشته باشد. فروشگاه‌ها ممکن است بیش از یک هدیه در لیست شما را پوشش دهند و هرگز کالاهایی که در انبار دارند تمام نمی‌شود. خانه شما فروشگاه $s_۱$ است که هیچ کالایی در انبار ندارد.

اقداماتی که در نظر خواهید گرفت، اقدامات سفر و خرید هستند که در آن ربات از مکان فعلی خود s_i به فروشگاه دیگری s_j در سریع‌ترین زمان ممکن سفر می‌کند و هر کالای باقی‌مانده در لیست خرید را که در s_j فروخته می‌شود، خریداری می‌کند. زمان سفر و خرید از s_i به s_j برابر با $t(s_i, s_j)$ است. شما می‌توانید فرض کنید که تمام اقدامات سفر و خرید کوتاه‌ترین مسیرها را نشان می‌دهند، بنابراین هیچ راه سریع‌تری برای رفتن بین s_i و s_j از طریق فروشگاه دیگری وجود ندارد. ربات در ابتدا در خانه شما بدون هیچ هدیه خریداری شده قرار دارد. شما می‌خواهید که ربات تمام اقلام را در کوتاه‌ترین زمان ممکن خریداری کند و به خانه بازگردد.

برای این مسئله برنامه‌ریزی، شما از یک فضای حالت استفاده می‌کنید که در آن هر حالت یک زوج (s, u) است که در آن s مکان فعلی است و u مجموعه هدایای خریداری نشده در لیست شما است (بنابراین $g \in u$ نشان می‌دهد که هدیه g هنوز خریداری نشده است).

(آ) اندازه فضای حالت بر حسب مقادیر تعریف شده در بالا چقدر است؟

(ب) برای هر یک از Heuristic های زیر که بر حالات (s, u) اعمال می‌شوند، مشخص کنید که آیا admissible، consistent، هیچ‌کدام یا هر دو هستند. فرض کنید که کمینه یک مجموعه خالی صفر است.

- کوتاهترین زمان از مکان فعلی به هر فروشگاه دیگر: $\min_{s' \neq s} t(s, s')$
- زمان رسیدن به خانه از مکان فعلی: $t(s, s_۱)$
- کوتاهترین زمان برای رسیدن به هر فروشگاهی که هر هدیه خریداری نشده را می‌فروشد:

$$\min_{g \in u}(\min_{s' : g \in S'} t(s, s'))$$

- کوتاهترین زمان برای رسیدن به خانه از هر فروشگاهی که هر هدیه خریداری نشده را می‌فروشد:

$$\min_{g \in u}(\min_{s' : g \in S'} t(s', s_۱))$$

(ج) حالتی را در نظر بگیرید که به جای ۱ ربات، از R ربات استفاده می‌کنیم که با سرعت برابر حرکت میکنند، حالا در حالت ابتدایی همه ربات ها در خانه هستند و در انتها دوباره به خانه برمی‌گردند و هر هدیه حداقل توسط یک ربات خریده شده است. فضای حالت مینیمال این مسئله را مشخص کنید. (راهنمایی: ربات ها لزوما همزمان به مقصد نمی‌رسند.)

پاسخ

(آ) اندازه فضای حالت برابر با $M \times ۲^N$ است. هر حالت یک زوج (s, u) است که در آن s مکان فعلی ربات است (یکی از M فروشگاه) و u زیرمجموعه‌ای از N هدیه است که هنوز خریداری نشده‌اند. چون برای هر هدیه دو حالت (خریداری شده یا نشده) وجود دارد، ۲^N حالت ممکن برای مجموعه هدایای خریداری نشده وجود دارد.

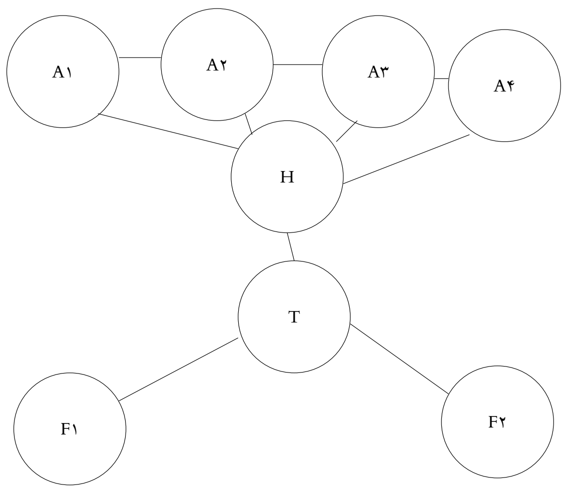
(ب)
- **کوتاهترین زمان از مکان فعلی به هر فروشگاه دیگر:** $\min_{s' \neq s} t(s, s')$
هیچ‌کدام. این هیوریستیک قابل قبول (admissible) نیست زیرا در حالت هدف (زمانی که ربات در خانه s_1 است و تمام هدایا خریداری شده‌اند)، مقدار آن صفر نمی‌شود. طبق تعریف، هیوریستیک باید در حالت هدف برابر با صفر باشد.
- **زمان رسیدن به خانه از مکان فعلی:** $t(s, s_1)$
هر دو (هم admissible و هم consistent). این هیوریستیک قابل قبول است زیرا ربات در نهایت باید به خانه بازگردد و $t(s, s_1)$ هزینه واقعی این بخش از مسیر است، بنابراین هرگز هزینه باقیمانده را بیش از حد تخمین نمی‌زند. همچنین سازگار (consistent) است زیرا هزینه سفر بین دو فروشگاه همواره بیشتر یا مساوی کاهش مقدار هیوریستیک است (نامساوی مثلثی).
- **کوتاهترین زمان برای رسیدن به هر فروشگاهی که هر هدیه خریداری نشده را می‌فروشد:** $\min_{g \in u}(\min_{s' : g \in s'} t(s, s'))$
هر دو (هم admissible و هم consistent). این هیوریستیک قابل قبول است زیرا ربات باید حداقل به یک فروشگاه دیگر برای خرید یک هدیه باقی‌مانده برود و این فرمول، کمترین هزینه ممکن برای این کار را در نظر می‌گیرد، بنابراین هزینه واقعی را کمتر تخمین می‌زند. این هیوریستیک سازگار نیز هست زیرا با حرکت از یک حالت به حالت دیگر، مقدار آن به اندازه‌ای بیشتر از هزینه واقعی حرکت کاهش نمی‌یابد.
- **کوتاهترین زمان برای رسیدن به خانه از هر فروشگاهی که هر هدیه خریداری نشده را می‌فروشد:** $\min_{g \in u}(\min_{s' : g \in s'} t(s', s_1))$
فقط admissible. این هیوریستیک قابل قبول است زیرا ربات باید پس از خرید آخرین هدیه به خانه برگردد و این فرمول کمترین هزینه ممکن برای این بازگشت را در نظر می‌گیرد، در نتیجه هزینه واقعی را کمتر تخمین می‌زند. اما سازگار نیست. برای مثال، فرض کنید از فروشگاه s_3 تا s_2 (که آخرین هدیه را دارد) هزینه ۱ و از s_2 تا خانه (s_1) هزینه ۵ باشد. در حالت (s_3, u)، مقدار هیوریستیک ۵ است. پس از حرکت به s_2 و خرید هدیه، حالت بعدی $(s_2, \emptyset)$ است و مقدار هیوریستیک ۰ می‌شود. در اینجا هزینه حرکت ۱ بود، اما مقدار هیوریستیک ۵ واحد کاهش یافت که این نقض شرط سازگاری است.

(ج) یک حالت باید مکان هر یک از R ربات و مجموعه هدایای خریداری نشده را مشخص کند. با توجه به اینکه ربات‌ها ممکن است بین فروشگاه‌ها در حال حرکت باشند، مکان هر ربات می‌تواند یکی از M فروشگاه یا یکی از مکان‌های بین راه باشد. اگر T حداکثر زمان سفر بین دو فروشگاه باشد، می‌توان برای هر ربات $M \times T$ مکان در نظر گرفت. همچنین ۲^N حالت برای هدایای خریداری نشده وجود دارد. بنابراین، اندازه فضای حالت برابر است با:

$$(MT)^R \times ۲^N$$

پرسش ۳ (۱۵ نمره)

گراف زیر را در نظر بگیرید.



(آ) برای گراف با استفاده از الگوریتم backtracking مسئله سه رنگ پذیری را حل کنید و مراحل را توضیح دهید.
اولویت ها از چپ به راست به شکل زیر هستند:

$$A۱, H, A۴, F۱, A۲, F۲, A۳, T \qquad R, G, B$$

(ب) همین سوال را این بار با MRV و Degree Heuristic (مرتبه تخمین و حداقل ارزش های باقی مانده) حل کنید.

(ج) توضیح دهید چرا انتخاب متغیر با بیشترین تعداد محدودیت و متغیر با کمترین تعداد مقدار باقی مانده اولویت بیشتری به بقیه دارند.

پاسخ (الف)

(آ) $A_1 = R$
(ب) $H = R$ با A_1 در تضاد است.
(ج) $H = G$
(د) $A_4 = R$
(ه) $F_1 = R$

(و) $A_{۲} = R$ با $A_{۱}$ در تضاد است، $A_{۲} = G$ با H در تضاد است، بنابراین $A_{۲} = B$.

(ز) $F_{۲} = R$

(ح) $A_{۳} = R$ با $A_{۴}$ در تضاد است، $A_{۳} = G$ با H در تضاد است، $A_{۳} = B$ با $A_{۲}$ در تضاد است، بنابراین **بازگشت به عقب** انجام می‌شود. مجموعه‌ی تعارض $\{A_{۲}, H, A_{۴}\}$ است، بنابراین پرش به $A_{۲}$ انجام می‌شود. $\{H, A_{۴}\}$ به مجموعه‌ی تعارض $A_{۲}$ اضافه می‌شود.

(ط) $A_{۲}$ مقدار دیگری ندارد، بنابراین بازگشت به عقب انجام می‌شود. مجموعه‌ی تعارض $\{A_{۱}, H, A_{۴}\}$ است، بنابراین پرش به $A_{۴}$ انجام می‌شود. $\{A_{۱}, H\}$ به مجموعه‌ی تعارض $A_{۴}$ اضافه می‌شود.

(ی) $A_{۴} = G$ با H در تضاد است، بنابراین $A_{۴} = B$.

(ک) $F_{۱} = R$

(ل) $A_{۲} = R$ با $A_{۱}$ در تضاد است، $A_{۲} = G$ با H در تضاد است، بنابراین $A_{۲} = B$.

(م) $F_{۲} = R$

(ن) $A_{۳} = R$

(س) $T = R$ با $F_{۱}$ و $F_{۲}$ در تضاد است، $T = G$ با H در تضاد است، بنابراین $T = B$.

(ع) **موفقیت**.

(ب)

(آ) $H = R$

(ب) $A_{۲} = R$ با H در تضاد است پس $A_{۲} = G$

(ج) $A_{۳} = R$ با H در تضاد است و $A_{۳} = G$ با $A_{۲}$ در تضاد است بنابراین $A_{۳} = B$

(د) $A_{۱} = R$ با H در تضاد است و $A_{۱} = G$ با $A_{۲}$ در تضاد است بنابراین $A_{۱} = B$

(ه) $A_{۴} = R$ با H در تضاد است. پس $A_{۴} = G$

(و) $T = R$ با H در تضاد است پس $T = G$

(ز) $F_{۱} = R$

(ح) $F_{۲} = R$

(ط) **موفقیت**.

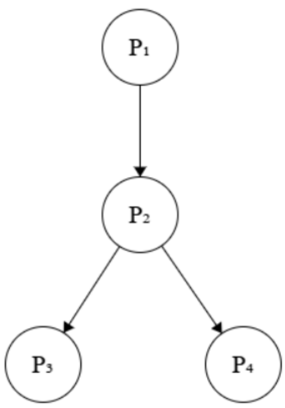
(ج) متغیری که بیشترین محدودیت را دارد منطقی است، زیرا متغیری را انتخاب می‌کند که (در شرایط برابر) احتمال بیشتری دارد که باعث شکست شود. شکست زودهنگام کارآمدتر است، زیرا بخش‌های بزرگی از فضای جستجو را هرس می‌کند.

روش مقدارگذاری با کمترین محدودیت نیز منطقی است، زیرا بیشترین فرصت را برای تخصیص‌های آینده فراهم می‌کند تا از وقوع تعارض جلوگیری شود.

پرسش ۴ (۲۰ نمره)

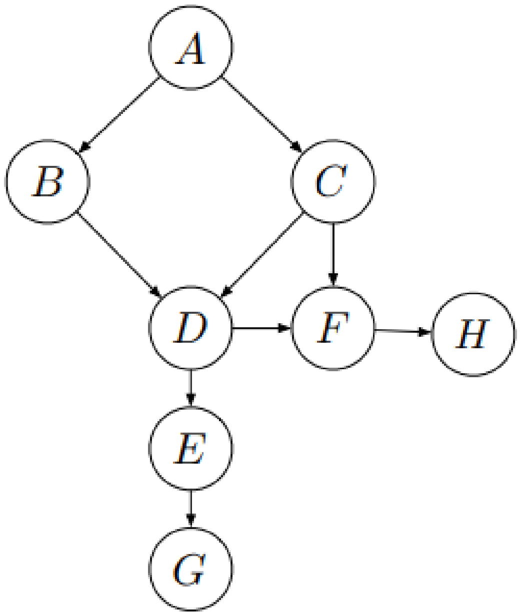
بخش اول با توجه به گراف زیر، و با استفاده از روش Inference by enumeration مقدار احتمال های زیر را حساب کنید.

$Pr(\neg p_3), Pr(p_2|\neg p_3), Pr(p_1|p_2, \neg p_3)$ a $Pr(p_1|\neg p_3, p_4)$



$$\Pr(p_{۱}) = ۰/۴$$
$$\Pr(p_{۲} \mid p_{۱}) = ۰/۸$$
$$\Pr(p_{۲} \mid \neg p_{۱}) = ۰/۵$$
$$\Pr(p_{۳} \mid p_{۲}) = ۰/۲$$
$$\Pr(p_{۳} \mid \neg p_{۲}) = ۰/۳$$
$$\Pr(p_{۴} \mid p_{۲}) = ۰/۸$$
$$\Pr(p_{۴} \mid \neg p_{۲}) = ۰/۵$$

بخش دوم با توجه به بیزنت زیر بدون در نظر گرفتن هیچ فرض اضافی دربارهٔ استقلال شرطی، به‌جز آن‌هایی که توسط نمودار بیان شده‌اند، برای هر مورد باذکر دلیل مشخص کنید که عبارت گفته شده تضمین میشود که برقرار است یا تضمین می شود که بر قرار نیست یا چیزی نمی توان گفت.



- $B \perp\!\!\!\perp C$
- $B \perp\!\!\!\perp C \mid G$
- $B \perp\!\!\!\perp C \mid H$
- $A \perp\!\!\!\perp D \mid G$
- $A \perp\!\!\!\perp D \mid H$
- $B \perp\!\!\!\perp C \mid A, F$
- $F \perp\!\!\!\perp B \mid D, A$
- $F \perp\!\!\!\perp B \mid D, C$

پاسخ الف)

$Pr(\neg p_3)$

$$\begin{aligned}
Pr(\neg p_۳) &= \sum_{P_۱,P_۲,P_۴} Pr(P_۱,P_۲,\neg p_۳,P_۴) = \sum_{P_۱,P_۲,P_۴} Pr(P_۱)Pr(P_۲|P_۱)Pr(\neg p_۳|P_۲)Pr(P_۴|P_۲) \\
&= Pr(p_۱)Pr(p_۲|p_۱)Pr(\neg p_۳|p_۲)Pr(p_۴|p_۲) + Pr(p_۱)Pr(p_۲|p_۱)Pr(\neg p_۳|p_۲)Pr(\neg p_۴|p_۲)+ \\
&\quad + Pr(p_۱)Pr(\neg p_۲|p_۱)Pr(\neg p_۳|\neg p_۲)Pr(p_۴|\neg p_۲) + Pr(p_۱)Pr(\neg p_۲|p_۱)Pr(\neg p_۳|\neg p_۲)Pr(\neg p_۴|\neg p_۲)+ \\
&\quad + Pr(\neg p_۱)Pr(p_۲|\neg p_۱)Pr(\neg p_۳|p_۲)Pr(p_۴|p_۲) + Pr(\neg p_۱)Pr(p_۲|\neg p_۱)Pr(\neg p_۳|p_۲)Pr(\neg p_۴|p_۲)+ \\
&\quad + Pr(\neg p_۱)Pr(\neg p_۲|\neg p_۱)Pr(\neg p_۳|\neg p_۲)Pr(p_۴|\neg p_۲) + Pr(\neg p_۱)Pr(\neg p_۲|\neg p_۱)Pr(\neg p_۳|\neg p_۲)Pr(\neg p_۴|\neg p_۲) \\
&= /۴ \times /۸ \times /۸ \times /۸ + /۴ \times /۸ \times /۸ \times /۲ + /۴ \times /۲ \times /۷ \times /۵ + /۴ \times /۲ \times /۷ \times /۵ + \\
&\quad + /۶ \times /۵ \times /۸ \times /۸ + /۶ \times /۵ \times /۸ \times /۲ + /۶ \times /۵ \times /۷ \times /۵ + /۶ \times /۵ \times /۷ \times /۵ \\
&= /۲۰۴۸ + /۰۵۱۲ + /۰۲۸ + /۰۲۸ + /۱۹۲ + /۰۴۸ + /۱۰۵ + /۱۰۵ = /۷۶۲
\end{aligned}$$

$Pr(p_۲|\neg p_۳)$

:$Pr(p_۲, \neg p_۳)$

$$\begin{aligned}
Pr(p_۲, \neg p_۳) &= \sum_{P_۱,P_۴} Pr(P_۱,p_۲,\neg p_۳,P_۴) = \sum_{P_۱,P_۴} Pr(P_۱)Pr(p_۲|P_۱)Pr(\neg p_۳|p_۲)Pr(P_۴|p_۲) \\
&= Pr(p_۱)Pr(p_۲|p_۱)Pr(\neg p_۳|p_۲)Pr(p_۴|p_۲) + Pr(p_۱)Pr(p_۲|p_۱)Pr(\neg p_۳|p_۲)Pr(\neg p_۴|p_۲)+ \\
&\quad + Pr(\neg p_۱)Pr(p_۲|\neg p_۱)Pr(\neg p_۳|p_۲)Pr(p_۴|p_۲) + Pr(\neg p_۱)Pr(p_۲|\neg p_۱)Pr(\neg p_۳|p_۲)Pr(\neg p_۴|p_۲) \\
&= /۴ \times /۸ \times /۸ \times /۸ + /۴ \times /۸ \times /۸ \times /۲ + /۶ \times /۵ \times /۸ \times /۸ + /۶ \times /۵ \times /۸ \times /۲ \\
&= /۲۰۴۸ + /۰۵۱۲ + /۱۹۲ + /۰۴۸ = /۴۹۶
\end{aligned}$$

$$Pr(p_۲|\neg p_۳) = \frac{Pr(p_۲, \neg p_۳)}{Pr(\neg p_۳)} = \frac{/۴۹۶}{/۷۶۲} = /۶۵۰۹$$

$Pr(p_۱|p_۲, \neg p_۳)$

$:Pr(p_۱, p_۲, \neg p_۳)$

$$\begin{aligned} Pr(p_۱, p_۲, \neg p_۳) &= \sum_{P_۴} Pr(p_۱, p_۲, \neg p_۳, P_۴) = \sum_{P_۴} Pr(p_۱)Pr(p_۲|p_۱)Pr(\neg p_۳|p_۲)Pr(P_۴|p_۲) \\ &= Pr(p_۱)Pr(p_۲|p_۱)Pr(\neg p_۳|p_۲)Pr(p_۴|p_۲) + Pr(p_۱)Pr(p_۲|p_۱)Pr(\neg p_۳|p_۲)Pr(\neg p_۴|p_۲) \\ &= /۴ \times /۸ \times /۸ \times /۸ + /۴ \times /۸ \times /۸ \times /۲ = /۲۰۴۸ + /۰۵۱۲ = /۲۵۶ \end{aligned}$$

$.Pr(p_۲, \neg p_۳) = /۴۹۶$

$$Pr(p_۱|p_۲, \neg p_۳) = \frac{Pr(p_۱, p_۲, \neg p_۳)}{Pr(p_۲, \neg p_۳)} = \frac{/۲۵۶}{/۴۹۶} = /۵۱۶۱$$

$Pr(p_۱|\neg p_۳, p_۴)$

$:Pr(p_۱, \neg p_۳, p_۴)$

$$\begin{aligned} Pr(p_۱, \neg p_۳, p_۴) &= \sum_{P_۲} Pr(p_۱, P_۲, \neg p_۳, p_۴) = \sum_{P_۲} Pr(p_۱)Pr(P_۲|p_۱)Pr(\neg p_۳|P_۲)Pr(p_۴|P_۲) \\ &= Pr(p_۱)Pr(p_۲|p_۱)Pr(\neg p_۳|p_۲)Pr(p_۴|p_۲) + Pr(p_۱)Pr(\neg p_۲|p_۱)Pr(\neg p_۳|\neg p_۲)Pr(p_۴|\neg p_۲) \\ &= /۴ \times /۸ \times /۸ \times /۸ + /۴ \times /۲ \times /۷ \times /۵ = /۲۰۴۸ + /۰۲۸ = /۲۳۲۸ \end{aligned}$$

$:Pr(\neg p_۳, p_۴)$

$$\begin{aligned} Pr(\neg p_۱, \neg p_۳, p_۴) &= \sum_{P_۲} Pr(\neg p_۱, P_۲, \neg p_۳, p_۴) = \sum_{P_۲} Pr(\neg p_۱)Pr(P_۲|\neg p_۱)Pr(\neg p_۳|P_۲)Pr(p_۴|P_۲) \\ &= Pr(\neg p_۱)Pr(p_۲|\neg p_۱)Pr(\neg p_۳|p_۲)Pr(p_۴|p_۲) + Pr(\neg p_۱)Pr(\neg p_۲|\neg p_۱)Pr(\neg p_۳|\neg p_۲)Pr(p_۴|\neg p_۲) \\ &= /۶ \times /۵ \times /۸ \times /۸ + /۶ \times /۵ \times /۷ \times /۵ = /۱۹۲ + /۱۰۵ = /۲۹۷ \end{aligned}$$

$$Pr(\neg p_۳, p_۴) = Pr(p_۱, \neg p_۳, p_۴) + Pr(\neg p_۱, \neg p_۳, p_۴) = /۲۳۲۸ + /۲۹۷ = /۵۲۹۸$$

$$Pr(p_۱|\neg p_۳, p_۴) = \frac{Pr(p_۱, \neg p_۳, p_۴)}{Pr(\neg p_۳, p_۴)} = \frac{/۲۳۲۸}{/۵۲۹۸} = /۴۳۹۴$$

ب) تنها مورد آخر قطعا درست است و بقیه موارد قابل تایید نیستند.

پرسش ۵ (۱۵ نمره)

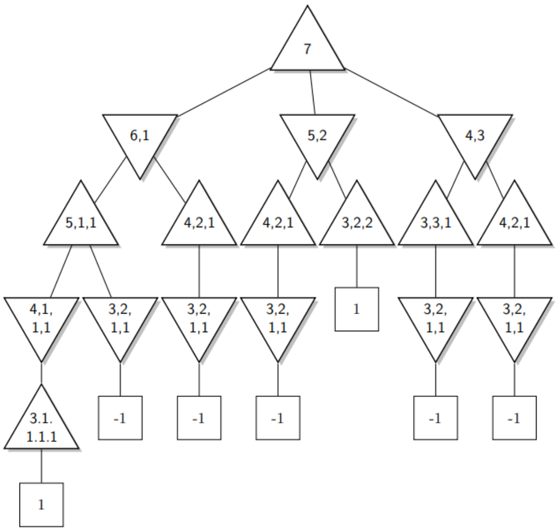
بازی دو نفره ای را در نظر بگیرید که با یک دسته ۷ تایی از قطعه های چوبی شروع می‌شود، در هر مرحله، فردی که نوبتش است باید یک دسته را به دو دسته تقسیم کند، به طوری که دو دسته برابر نباشند. اگر همه دسته ها از یک یا دو قطعه تشکیل شده باشند، بازی تمام می شود فردی که نوبت او بوده می‌بازد.

(آ) درخت این بازی را رسم کنید.

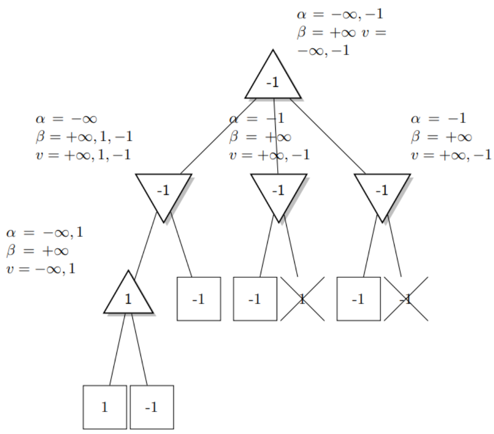
(ب) فرض کنید به ازای استراتژی هایی که در آنها بازیکن max برنده است در برگ مقدار یک و در غیر اینصورت مقدار منفی یک گذاشته ایم. الگوریتم minimax را به همراه alpha-beta pruning برای این بازیکن اجرا کنید تا بهترین حرکت برای بازیکن max در ریشه پیداکنید. (جهت انجام alpha-beta pruning را از چپ به راست در نظر بگیرید)

پاسخ

الف)



ب)



پرسش ۶ (۱۸ نمره)

مدل HMM ای را درنظر بگیرید که متغیر مخفی آن، آفتابی (C) یا ابری (S) بودن هوا است. احتمال اولیه $P(X_1 = S) = 0.8$ و احتمالات انتقال $P(X_t = S|X_{t-1} = S) = 0.1$ و $P(X_t = C|X_{t-1} = C) = 0.1$ است. شما هر روز همسایه خود را مشاهده می‌کنید که به خانه برمی‌گردد، او هر روز دقیقا یک مورد از بین تیشرت، شلوارک و هودی را انتخاب کرده است. اگر جدول احتمالات مشاهدات به شکل زیر باشد، و شما روز اول شلوارک و روز بعد تیشرت را مشاهده کرده باشید، به سوالات زیر پاسخ دهید..

E_t	t-shirt	short	hoodie
$P(E_t\|X_t = S)$	۰/۴	۰/۴	۰/۲
$P(E_t\|X_t = C)$	۰/۲	۰/۲	۰/۶

(آ) تمام ترکیبات ممکن X_1, X_2 را مشخص کنید و مقدار $P(X_1, E_1, X_2, E_2)$ را برای آنها حساب کنید.

(ب) مقدار احتمال $P(X_2 = C|E_1 = short, E_2 = t - shirt)$

(ج) از الگوریتم viterbi برای پیدا کردن محتمل ترین وضعیت هوا با توجه به مشاهدات دو روز استفاده کنید.

پاسخ

(آ) با توجه به اینکه $P(X_۱ = S) = ۰/۸$ و $P(X_۱ = C) = ۰/۲$ و همچنین $P(X_t = C|X_{t-۱} = S) = ۰/۹$ و $P(X_t = S|X_{t-۱} = C) = ۰/۹$ و با استفاده از فرمول $P(X_۱, E_۱, X_۲, E_۲) = P(X_۱)P(E_۱|X_۱)P(X_۲|X_۱)P(E_۲|X_۲)$، احتمالات به صورت زیر محاسبه می‌شوند:

- $P(X_۱ = S, E_۱ = \text{short}, X_۲ = S, E_۲ = \text{t-shirt}) = (۰/۸)(۰/۴)(۰/۱)(۰/۴) = ۰/۰۱۲۸$
- $P(X_۱ = S, E_۱ = \text{short}, X_۲ = C, E_۲ = \text{t-shirt}) = (۰/۸)(۰/۴)(۰/۹)(۰/۲) = ۰/۰۵۷۶$
- $P(X_۱ = C, E_۱ = \text{short}, X_۲ = S, E_۲ = \text{t-shirt}) = (۰/۲)(۰/۲)(۰/۹)(۰/۴) = ۰/۰۱۴۴$

- $P(X_1 = C, E_1 = \text{short}, X_2 = C, E_2 = \text{t-shirt}) = (0.2)(0.2)(0.1)(0.2) = 0.0008$

(ب) برای محاسبه این احتمال شرطی، از الگوریتم پیشرو (forward algorithm) و مقادیر محاسبه شده در قسمت قبل استفاده می‌کنیم. ابتدا احتمال مشترک $P(E, X_2 = C)$ را با جمع کردن تمام مسیرهایی که به $X_2 = C$ ختم می‌شوند، به دست می‌آوریم:

$$P(E, X_2 = C) = P(X_1 = S, E, X_2 = C) + P(X_1 = C, E, X_2 = C)$$

$$= 0.0576 + 0.0008 = 0.0584$$

سپس احتمال کل مشاهدات $P(E)$ را با جمع تمام احتمالات از قسمت (الف) محاسبه می‌کنیم:

$$P(E) = 0.0128 + 0.0576 + 0.0144 + 0.0008 = 0.0856$$

در نهایت، احتمال شرطی برابر است با:

$$P(X_2 = C|E) = \frac{P(E, X_2 = C)}{P(E)} = \frac{0.0584}{0.0856} \approx 0.6822$$

(ج) الگوریتم Viterbi محتمل‌ترین توالی حالات را پیدا می‌کند. این کار با مقایسه احتمالات مسیرهای مختلف انجام می‌شود.

- برای یافتن بهترین حالت قبلی برای $X_2 = S$، دو احتمال زیر را مقایسه می‌کنیم:
  - $P(X_1 = S, E_1, X_2 = S, E_2) = 0.0128$
  - $P(X_1 = C, E_1, X_2 = S, E_2) = 0.0144$

  مقدار بزرگتر 0.0144 است، بنابراین محتمل‌ترین حالت قبلی برای $X_2 = S$، حالت $X_1 = C$ است.
- برای یافتن بهترین حالت قبلی برای $X_2 = C$، دو احتمال زیر را مقایسه می‌کنیم:
  - $P(X_1 = S, E_1, X_2 = C, E_2) = 0.0576$
  - $P(X_1 = C, E_1, X_2 = C, E_2) = 0.0008$

  مقدار بزرگتر 0.0576 است، بنابراین محتمل‌ترین حالت قبلی برای $X_2 = C$، حالت $X_1 = S$ است.
- برای پیدا کردن بهترین حالت نهایی، مقادیر ماکزیمم هر ستون را با هم مقایسه می‌کنیم:

  $$\max(0.0144, 0.0576) = 0.0576$$

  این مقدار مربوط به حالتی است که $X_2 = C$ بوده و از حالت $X_1 = S$ آمده است.

بنابراین، محتمل‌ترین توالی وضعیت هوا $(X_1, X_2) = (S, C)$ یا (آفتابی، ابری) است.