



UNIVERSITY  
OF WOLLONGONG  
AUSTRALIA

University of Wollongong  
Research Online

---

Faculty of Commerce - Papers (Archive)

Faculty of Business

---

2005

# Application of human reasoning in question answering systems

Farhad Oroumchian

*University of Wollongong*, farhado@uow.edu.au

E. Darudi

M. Rahgozar

B. Ofogi

---

## Publication Details

Oroumchian, F., Darudi, E., Rahgozar, M. & Ofogi, B. (2005). Application of human reasoning in question answering systems. In A. Khademzadeh, S. Shahnazi & K. Badie (Eds.), *Proceedings of 10th Annual Computer Society of Iran Computer Conference CSCIC 2005* (pp. 373-380). Tehran, Iran: Sadegh.

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library:  
research-pubs@uow.edu.au

---

# Application of human reasoning in question answering systems

**Keywords**

Application, Human, Reasoning, Question, Answering, Systems

**Disciplines**

Business | Social and Behavioral Sciences

**Publication Details**

Oroumchian, F., Darudi, E., Rahgozar, M. & Ofogi, B. (2005). Application of human reasoning in question answering systems. In A. Khademzadeh, S. Shahnazi & K. Badie (Eds.), Proceedings of 10th Annual Computer Society of Iran Computer Conference CSCIC 2005 (pp. 373-380). Tehran, Iran: Sadegh.

## کاربرد استدلال‌های انسانی در سامانه‌های جوابگویی پرسش

احسان درودی<sup>۱</sup>، مسعود رهگذر<sup>۲</sup>، فرهاد ارومچیان<sup>۳</sup> و بهادر رضا افقی<sup>۴</sup>

<sup>۱</sup> مرکز تحقیقات مخابرات ایران

<sup>۲</sup> گروه مهندسی برق و کامپیوتر، دانشکده فنی دانشگاه تهران

<sup>۳</sup> قطب علمی کنترل و پردازش هوشمند، گروه مهندسی برق و کامپیوتر، دانشکده فنی دانشگاه تهران

<sup>۴</sup> University of Wollongong in Dubai

### چکیده

سامانه‌های جوابگویی پرسش<sup>۱</sup> سعی دارند تا جواب‌های دقیقی برای سوالهایی که به زبان طبیعی پرسیده میشوند پیدا کنند. در این مقاله سامانه‌ی جوابگویی پرسشی بر مبنای تئوری استدلال مقبول انسانی<sup>۲</sup> تشریح میگردد. ایده اصلی در این تحقیق، بهره‌گیری از منطق استدلال مقبول در موتور استنباطی این سامانه برای پیدا کردن جوابهای قابل قبول است، هنگامی که جوابهای صریح و دقیق در دسترس نباشند. زیرمجموعه‌ای از استنباط‌های تئوری استدلال مقبول انتخاب شده و الگوریتمی برای انجام استدلال زنجیره‌ای با این استنباط‌های پایه پیشنهاد و پیاده‌سازی گردیده است. آزمایشهای صورت گرفته با زیرمجموعه‌ای از پایگاه دانش TeLQAS نشانگر نتایج امیدوارکننده‌ای بوده است.

### کلمات کلیدی

استدلال مقبول انسانی، سامانه جوابگویی پرسش، بازیابی اطلاعات، سامانه‌ی خبره

### ۱- مقدمه

حجم دانش بشری ذخیره شده بصورت الکترونیکی در حال افزایش است و این مساله ساخت ابزارهایی را برای دسترسی به این حجم عظیم اطلاعات ضروری می‌سازد. ساده‌ترین این ابزارها موتورهای بازیابی اطلاعات<sup>۳</sup> هستند که با وارد کردن چند کلمه-ی کلیدی، تعدادی سند را باز میگردانند. هدف از ایجاد سامانه‌های جوابگویی پرسش ساده‌تر کردن نحوه‌ی ارتباط با کاربر است. به این ترتیب که کاربر پرسش خود را در قالب زبان طبیعی بیان کرده و سامانه هم سعی می‌کند تا جواب دقیق را استخراج کرده و ارائه نماید. بدینگونه، وظیفه‌ی نوشتن کلمات کلیدی از دوش کاربر برداشته میشود و از سوی دیگر لازم نخواهد بود که وی برای پیدا کردن جواب، مستندات برگشتی را مطالعه کند.

<sup>۱</sup> Question Answering Systems

<sup>۲</sup> Human Plausible Reasoning

<sup>۳</sup> Information Retrieval

در این مقاله سامانه‌ی جوابگویی پرسش پیشنهادی بر مبنای تئوری استدلال مقبول انسانی [۱] تشریح می‌گردد. این تئوری بر مبنای مشاهده‌ی رفتار انسانها در محاوره‌های واقعی جوابگویی به پرسش شکل گرفته است، هنگامی که اطلاعات لازم برای جوابگویی، ناقص و یا با عدم قطعیت همراه است. یک مکانیزم استدلال مشابه آنچه انسانها بکار می‌برند میتواند از اطلاعات ضمنی، ناقص و نامطمئن موجود استفاده کرده و کارایی سامانه‌ی جوابگویی پرسش را بهبود بخشد؛ چراکه این سامانه‌ها هم در ورودی و هم در خروجی با انسان و مفاهیم بیان شده با زبان انسان مواجه هستند. سامانه‌ی پیشنهادی، مانند دیگر سامانه‌های جوابگویی پرسش، سعی میکند تا جواب دقیقی به پرسش وارد شده بدهد، ولی در صورتی که این کار ممکن نباشد، متمایز از دیگر سامانه‌ها، تلاش می‌کند تا یک جواب قابل قبول بیاید.

در ادامه‌ی این مقاله، ابتدا در بخش دوم مقاله تحقیقات مرتبط را بصورت اجمالی بررسی میکنیم. در بخش سوم، نگاهی دقیق‌تر به تئوری استدلال مقبول خواهیم داشت و سپس در بخش چهارم معماری سامانه‌ی جوابگویی پرسش ما بطور خلاصه ارائه می‌شود. در بخش پنجم، روش بازنمایی دانش برای استفاده در موتور استدلالی آمده است. مجموعه استنباط‌های انتخاب شده از تئوری استدلال مقبول برای پیاده‌سازی در سامانه در بخش ششم ارائه می‌گردد. الگوریتم استدلال پیشنهادی در بخش هفتم آمده و بخش هشتم، نتایج اولیه پیاده‌سازی برای اعتبارسنجی این الگوریتم را نشان میدهد. در نهایت در بخش نهم، نتیجه‌گیری و کارهای برنامه ریزی شده برای آینده ذکر میگردد.

## ۲- کارهای مربوطه

تئوری استدلال مقبول در گذشته با موفقیت در موارد کاربردی گوناگون مبتنی بر اطلاعات به خدمت گرفته شده است. در [۲] از یک پیاده‌سازی آزمایشی تئوری برای پاسخگویی به سوالهایی مربوط به جدول تناوبی عناصر شیمیایی استفاده شده است و در [۳] یک سامانه‌ی خیره برای تشخیص نوع چمن بر مبنای این تئوری پیاده سازی شده است. ارومچیان تئوری استدلال مقبول را برای ایجاد یک سامانه‌ی تجربی بازبایی اطلاعات با نام PLIR بکار گرفته است [۴]. در [۵]، [۶] و [۷] نویسندگان کاربردهایی از این تئوری را به ترتیب برای تصفیه تطبیقی اطلاعات، آموزش هوشمند و دسته‌بندی مستندات ارائه داده‌اند. این تئوری همچنین در یک واسط کاربر گرافیکی که به کاربران توصیه‌های هوشمندانه می‌کند بکار رفته است [۸]. این پیاده‌سازی‌ها بر مفید بودن و انعطاف پذیری بالای تئوری استدلال مقبول برای مواردی که نیاز به فعل و انفعال و استدلال در مورد انسانها دارند، صحنه می‌گذارند.

## ۳- تئوری استدلال مقبول انسانی

تئوری استدلال مقبول سعی در توصیف الگوهای استنباطی متفاوتی دارد که انسانها برای استدلال در مورد جهان پیرامون بکار می‌برند و طیف وسیعی از استنباط‌های را در بر میگیرد که انسانها در زندگی روزمره از آنها استفاده میکنند. در این تئوری، بیشتر دانش توسط بیانهای<sup>۱</sup> ساده نشان داده می‌شود. بعنوان مثال میتوان بیان زیر را در نظر گرفت؛

$$\text{Color (Eyes)} = \{\text{Blue}\}, \gamma = 1.0$$

در این بیان، Color یک توصیفگر<sup>۲</sup> است و Eyes و Blue به ترتیب آرگومان<sup>۳</sup> و مورد ارجاع<sup>۴</sup> نامیده می‌شوند.  $\gamma$  پارامتری است که میزان اطمینان به این بیان را نشان داده و علاوه بر  $\gamma$ ، پارامترهای دیگری هم در تئوری استدلال مقبول وجود دارند که تنها تعدادی از آنها برای پیاده‌سازی در این سامانه انتخاب شده‌اند. این پارامترها عبارتند از؛ درجه‌ی تسلط<sup>۵</sup> (۸)، درجه‌ی

<sup>۱</sup> Statement

<sup>۲</sup> Descriptor

<sup>۳</sup> Argument

<sup>۴</sup> Referent

<sup>۵</sup> Dominance

شبهات ( $\sigma$ ) و احتمال شرطی<sup>۱</sup> ( $\alpha$ ). پارامتر  $\delta$  میزان تسلط<sup>۲</sup> یک زیرمجموعه را در مجموعه‌ی مربوطه نشان می‌دهد. پارامتر  $\sigma$  میزان شبهات بین دو مفهوم را نشان داده و پارامتر  $\alpha$  نیز، همانطور که در ادامه توضیح داده شده، برای نشان دادن میزان وابستگی مفاهیم بکار می‌رود. شکل دیگر بازنمایی دانش، عبارت‌های وابستگی<sup>۳</sup> هستند. بعنوان مثال، وابستگی زیر بیان می‌کند که ما مطمئن هستیم که میزان متوسط درجه حرارت یک محل با احتمال زیاد وابسته به عرض جغرافیایی آن محل است؛

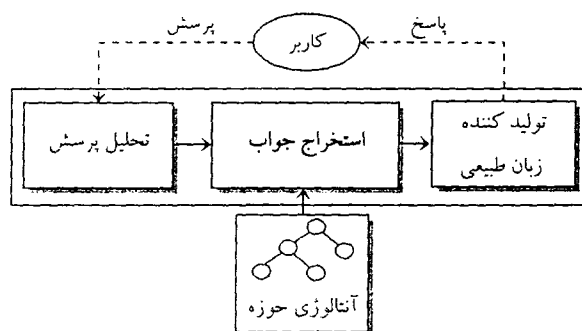
Latitude (Place)  $\rightarrow$  Average Temperature (Place),  $\alpha = \text{high}$ ,  $\gamma = \text{high}$

وابستگی‌ها می‌توانند علامت‌دار یا بی‌علامت باشند. یک وابستگی علامت‌دار می‌تواند مثبت یا منفی باشد. در وابستگی‌های مثبت (که در مثال‌ها با علامت  $\rightarrow$  نشان داده شده‌اند)، افزایش یک طرف وابستگی موجب افزایش مفهوم طرف دیگر می‌شود. این مطلب در مورد وابستگی‌های منفی برعکس می‌باشد. وابستگی بدون علامت، صرفاً<sup>۴</sup> بیان می‌کند که بین دو مفهوم یک وابستگی وجود دارد، بدون آنکه بخواهیم یا بتوانیم آنرا بصورت کمی بیان کنیم (مانند مثال بالا).

استنباط‌های تئوری استدلال مقبول روی این عبارت‌های منطقی شکل می‌گیرند. در بخش ششم، این استنباط‌ها بگونه‌ای که پیاده‌سازی شده‌اند توضیح داده خواهند شد.

#### ۴- معماری سامانه‌ی پرسش و پاسخ

TeLQAS [۹] یک سامانه‌ی جوابگویی پرسش حوزه-محدود<sup>۴</sup> و مبتنی بر آنتالوژی<sup>۵</sup> می‌باشد که برای حوزه‌ی مخابرات در مرکز تحقیقات مخابرات ایران طراحی و پیاده‌سازی شده است. بطور کلی، جریان داده در این سامانه از دو فرآیند برخط<sup>۶</sup> و برون‌خط<sup>۷</sup> تشکیل شده است که بخش برون‌خط وظیفه‌ی جمع دانش مورد نیاز سامانه را در قالب یک آنتالوژی را بر عهده دارد و بخش برخط در عمل به پرسش‌های کاربران بر اساس دانش موجود در آنتالوژی جواب می‌دهد. معماری کامل این سامانه پیچیده و شامل مولفه‌های زیادی است که خارج از حوصله‌ی این مقاله می‌باشد. در اینجا صرفاً<sup>۸</sup> قسمتی از بخش برخط که مرتبط با این تحقیق است بررسی می‌شود.



شکل ۱- جریان داده در بخش برخط سامانه‌ی TeLQAS

شکل (۱) معماری بخش برخط TeLQAS را نشان می‌دهد. در فرآیند برخط، ابتدا پرسش وارد شده به سامانه تحلیل و بررسی شده و سوال از زبان طبیعی به یک بازنمایی مناسب‌تر برای مولفه‌ی استخراج جواب تبدیل می‌شود که آنرا پرسش مقبول می‌

<sup>۱</sup> Conditional Likelihood

<sup>۲</sup> Dominance

<sup>۳</sup> Dependency

<sup>۴</sup> Restricted-Domain

<sup>۵</sup> Ontology

<sup>۶</sup> Online

<sup>۷</sup> Offline

نامیم. این پرسشهای مقبول بیان های ناقصی هستند که مولفهی استخراج جواب سعی می کند آنها را کامل کند. نمونه هایی از پرسش های زبان طبیعی به همراه پرسش های مقبول متناظر آنها در جدول (۱) آمده اند. مهمترین مولفهی فرآیند برخط، استخراج کننده ی جواب (موتور استدلال) میباشد که بر اساس پرسش تحلیل شده، جواب را از آنتالوژی استخراج نموده و به همراه توجیه جواب برای جمله بندی به مولفهی تولید کننده ی زبان طبیعی ارسال می کند. دانش مورد نیاز برای این موتور استدلال در سامانه بصورت خاصی بازنمایی میشود که در بخش بعدی توضیح داده شده است.

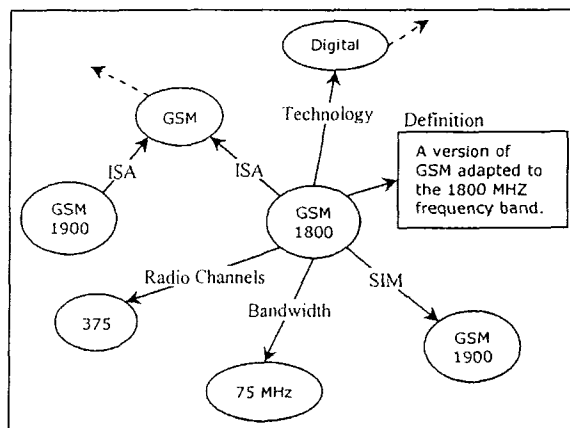
جدول ۱- نمونه هایی از تبدیل پرسش زبان طبیعی به شکل مقبول آنها توسط مولفهی تحلیل پرسش

پرسش مقبول	پرسش به زبان طبیعی
Definition ("fiber optic") = {?}	What is fiber optic?
Causes ("photon loss") = {?}	What causes photon loss?
Has_Type ("multiple access techniques") = {?}	What are different types of multiple access techniques?

## ۵- بازنمایی دانش

دانش سامانه ی TeLQAS در یک آنتالوژی نگهداری می شود که برای مولفهی استدلال بصورت یک شبکه ی معنایی<sup>۱</sup> پیاده سازی شده است. شکل (۲) گوشه ای از این شبکه معنایی را برای حوزه ی مخابرات سیار TeLQAS نشان می دهد. هر یک از گره ها در این شبکه ی معنایی، میتواند نماینده ی یک آرگومان یا مورد ارجاع باشد. روابط (پیوندهای) بین گره ها نیز نمایانگر توصیفگرها هستند. هر پیوند به همراه گره های دو سر پیوند، یک بیان منطقی را بازنمایی می کند. به عنوان مثال، در شکل (۲) میتوان بیان  $\{Digital\} = Technology(GSM\ 1800)$  را تشخیص داد.

بعضی از روابط در تئوری استدلال مقبول معنای خاصی دارند. رابطه ISA ساختارهای سلسله مراتبی لازم را برای استنباط های تعمیم و تخصیص مهیا کرده و رابطه ی SIM نیز برای استنباط شباهت بکار می رود (این استنباط ها در بخش ششم توضیح داده می شوند). وابستگی های منطقی مقبول می توانند نماینده ی هر رابطه ای از پایگاه دانش باشند که بر علیت، وابستگی، تاثیر گذاری و غیره دلالت میکند. استنباط های مبتنی بر وابستگی بر اساس این رابطه ها شکل میگیرند.



شکل ۲- نمایی از بازنمایی آنتالوژی TeLQAS با شبکه معنایی

<sup>۱</sup> Semantic Network

## ۶- استنباط‌های مقبول پایه

این بخش مجموعه‌ای از استنباط‌های تئوری استدلال مقبول را بگونه‌ای که پیاده‌سازی شده‌اند توضیح می‌دهد. این مجموعه شامل سه تبدیل<sup>۱</sup> و دو استنباط مبتنی بر وابستگی است. تبدیلات، ساده‌ترین شکل استنباط هستند و با استفاده از روابط سلسله مراتبی ISA و SIM قابل انجام می‌باشند. اگر بدانیم که یک بیان با درجه‌ای از اطمینان درست است، آنگاه می‌توانیم سه تبدیل را روی این بیان انجام داده و سه بیان جدید با درجه‌های مختلف اطمینان بدست آوریم. این تبدیلات عبارتند از: تعمیم آرگومان (AGEN)، تخصیص آرگومان (ASPEC) و شباهت آرگومان (ASIM).

شکل (۳) تبدیل تعمیم آرگومان را به همراه یک مثال از پایگاه دانش TeLQAS نشان می‌دهد. در این استنباط میزان اطمینان به جواب نهایی تابعی است از درجه‌ی اطمینان به بیان اولیه، اطمینان به رابطه‌ی GEN (همان ISA) و درجه تسلط زیرمجموعه در مجموعه‌ی اصلی. در حال حاضر این تابع (و دیگر توابع ترکیب) حاصلضرب معمولی این پارامترها می‌باشد.

Des (Arg <sub>1</sub> ) = {Ref}	: $\gamma_{stat}$
Arg <sub>2</sub> <u>SIM</u> Arg <sub>1</sub>	: $\gamma_{sim}, \sigma$
-----	
Des (Arg <sub>2</sub> ) = {Ref}	: $\gamma_{con} = f(\gamma_{stat}, \gamma_{sim}, \sigma)$
مثال:	
Modulation_Type ("GSM 1800") = {"GSMK"} "GSM 1900" <u>SIM</u> "GSM 1800"	
-----	
Modulation_Type ("GSM 1900") = {"GSMK"}	

شکل ۴- تبدیل شباهت آرگومان

Des (Arg <sub>1</sub> ) = {Ref}	: $\gamma_{stat}$
Arg <sub>2</sub> <u>GEN</u> Arg <sub>1</sub>	: $\gamma_{gen}, \delta$
-----	
Des (Arg <sub>2</sub> ) = {Ref}	: $\gamma_{con} = f(\gamma_{stat}, \gamma_{gen}, \delta)$
مثال:	
Technology ("GSM 1800") = {Digital} GSM <u>GEN</u> "GSM 1800"	
-----	
Technology (GSM) = {Digital}	

شکل ۳- تبدیل تعمیم آرگومان

استنباط تخصیص آرگومان شبیه استنباط تعمیم است، با این تفاوت که سلسله مراتب ISA را به سمت پایین پیمایش میکند. استنباط شباهت آرگومان، روابط SIM را به خدمت گرفته و سعی می‌کند تا سلسله مراتب را در عرض پیمایش کرده و با مقایسه، جواب را بیابد. شکل (۴) این استدلال را به همراه یک مثال نشان می‌دهد. میزان اطمینان به جواب نهایی تابعی است از درجه‌ی اطمینان به بیان اولیه، درجه‌ی اطمینان به رابطه‌ی SIM و درجه‌ی شباهت دو مفهوم.

استنباط‌های مبتنی بر وابستگی عبارتند از: اشتقاق از وابستگی<sup>۲</sup> و استنباط تعدی<sup>۳</sup>. اشتقاق از وابستگی، قابل اعمال به وابستگی‌های مثبت است (در حال حاضر در پایگاه دانش ما فقط وابستگی‌های مثبت وجود دارند). هنگامی که یک رابطه‌ی وابستگی بین دو مفهوم برقرار باشد، اگر مقدار مفهوم تاثیرگذار زیاد، متوسط یا کم باشد، آنگاه مفهوم تاثیرپذیر نیز بطور متناظر زیاد، متوسط یا کم خواهد بود (این مطلب در مورد وابستگی‌های منفی معکوس خواهد بود). شکل (۵) یک استنباط اشتقاق از وابستگی را نشان می‌دهد. میزان اطمینان به جواب نهایی تابعی از درجه‌ی اطمینان به بیان اولیه و رابطه‌ی وابستگی و میزان پارامتر احتمال شرطی رابطه‌ی وابستگی می‌باشد.

استنباط تعدی هنگامی که رابطه‌ی وابستگی مستقیمی بین دو مفهوم وجود نداشته باشد، سعی می‌کند تا با ترکیب دو رابطه‌ی وابستگی مرتبط، آن را تولید کند. بعنوان مثال اگر مولفه‌ی استدلال بداند که Fiber Alignment باعث Splice Loss می‌شود و Splice Loss خود روی Gain Deterioration تاثیر می‌گذارد، آنگاه نتیجه می‌گیرد که Fiber Alignment نیز، با درجه اطمینان و احتمال شرطی کوچک‌تر، مستقیماً روی Gain Deterioration تاثیر می‌گذارد.

<sup>۱</sup> Transform<sup>۲</sup> Derivation from Dependency (DDEP)<sup>۳</sup> Transitivity Inference

$Des_1(a) \rightarrow Des_2(a)$	$: \gamma_{dep, \alpha}$
$Des_1(Arg) = \{high \dots low\}$	$: \gamma_{stat}$
Arg SPEC a	$: \gamma_{spec}$
-----	
$Des_2(Arg) = \{high \dots low\}$	$: \gamma_{con} = f(\gamma_{dep}, \gamma_{stat}, \alpha, \gamma_{spec})$
مثال:	
Refractive Index (Coating) $\rightarrow$ Optical Loss (Coating)	
Refractive_Index (Cladding) = high	
Cladding ISA Coating	
-----	
Optical_Loss (Cladding) = high	

شکل ۵- استنباط اشتقاق از وابستگی

## ۷- الگوریتم استدلال

در حالت کلی، اعمال استنباط‌های پایه ممکن است به تنهایی منجر به جواب نشود. تئوری استدلال مقبول توضیح نداده است که انسانها چگونه استنباط‌های پایه را با هم ترکیب می‌کنند. در این بخش یک الگوریتم استدلال زنجیره‌ای را پیشنهاد می‌دهیم که در آن استنباط‌های پایه می‌توانند به هم متصل شوند، به نحوی که پیش‌شرط‌های هر استنباط، پرسش‌هایی برای استنباط‌های بعدی باشند. استدلال زمانی متوقف می‌شود که یک جواب صریح در پایگاه دانش پیدا شده یا اینکه شرایط محدود کننده برقرار شوند. بدین ترتیب این الگوریتم بصورت استدلال زنجیره‌ای رو به عقب<sup>۱</sup> عمل می‌کند.

موتور استدلالی ابتدا یک پرسش مقبول را از مولفه‌ی تحلیل پرسش دریافت کرده و بررسی می‌کند که آیا جواب صریح متناظر در پایگاه دانش وجود دارد یا خیر. اگر جواب صریح در دسترس باشد، سامانه آن را برمی‌گرداند و در غیر این صورت سعی می‌کند جواب قابل قبولی را با اعمال تمامی استنباط‌های پایه‌ی ممکن روی پرسش بدست آورد. هر استنباط پایه در ابتدا پرسش مقبول جدیدی برای خود ایجاد می‌کند تا روی آن کار کند (به عنوان مثال، یک استنباط تعمیم آرگومان پرسشی می‌سازد که در آن آرگومان نسبت به پرسش اولیه تعمیم داده شده است). سپس بررسی می‌شود که آیا جواب صریح به این پرسش جدید وجود دارد یا خیر. اگر جواب موجود بود، درجه اطمینان به آن ( $\gamma$ ) را محاسبه کرده و به همراه جواب به تابع استنباط مرجع (تابع استنباط صدا زنده) بازمی‌گرداند. اگر جواب صریح در پایگاه دانش موجود نباشد، استنباط خود سعی می‌کند با صدا زدن دیگر استنباط‌ها جواب قابل قبولی را بدست آورد و این کار بصورت بازگشتی ادامه پیدا می‌کند. در تمامی مراحل استدلال و پس از اینکه یک استنباط تمامی جواب‌های مقبول را بدست آورد، آنها را با هم ترکیب می‌کند تا بهترین جواب (جواب با بیشترین مقدار  $\gamma$ ) را به تابع استنباط مرجع برگرداند.

در الگوریتم استدلال، برای هر استنباط، اولویتی در نظر گرفته شده است و استنباط‌ها به ترتیب اولویت انجام می‌شوند. این اولویت‌بندی به این دلیل صورت گرفته که روند استدلال ابتدا در جهاتی هدایت شود که شانس پیدا کردن جواب بیشتر است. بنظر می‌رسد که استنباط‌های تعمیم و تخصیص بیش از همه توسط انسانها بکار گرفته می‌شوند. از اینرو به آنها اولویت بالاتری داده شده و زودتر از سایر استنباط‌ها اجرا می‌شوند. سپس استنباط‌های شباهت و اشتقاق از وابستگی به ترتیب اجرا می‌گردند. استنباط تعدی نیز صرفاً<sup>۲</sup> توسط استنباط اشتقاق از وابستگی در مواقع لزوم صدا زده می‌شود.

از طرف دیگر، انسانها معمولاً در یک زمان به موضوعات گوناگون فکر نمی‌کنند و معمولاً استدلال زنجیره‌ای را تا عمق زیادی گسترش نمی‌دهند، چراکه اطمینان به جواب بشدت افت کرده و دیگر قابل قبول نخواهد بود. به این دلیل الگوریتم فراخوانی استدلالها بصورت جستجوی عمق اول<sup>۲</sup> با عمق محدود پیاده‌سازی شده است. به این ترتیب، استنباط‌ها با اولویت ذکر

<sup>۱</sup> Backward Chaining Reasoning<sup>۲</sup> Depth First Search



شده در بالا، بصورت عمق-اول انجام شده و اگر عمق از عددی معین بیشتر شود، الگوریتم به عمق قبلی بازگشت<sup>۱</sup> می‌کند و کار را روی دیگر استنباطها ادامه می‌دهد.

## ۸- اعتبارسنجی الگوریتم استدلال

اعتبارسنجی الگوریتم استدلال برای سامانه‌ی پرسش و پاسخ TeLQAS روی بخش کوچکی از آنتالوژی این سامانه در حوزه مخابرات نوری انجام شد (شامل حدود ۵۰۰ مفهوم و ۶۰۰ رابطه). تعدادی سوال که معمولاً توسط افراد نیمه‌متخصص در این حوزه پرسیده می‌شود، انتخاب گردید و جوابهای سامانه به همراه توجیه آنها بررسی شد. جدول (۲) نتایج حاصل شده را نشان می‌دهد.

جدول ۲- چند پرسش انتخاب شده به همراه جوابها و توجیه‌های سامانه

پرسش	پرسش مقبول	جواب سامانه	نوع	درجه‌ی اطمینان	توجیه سامانه برای جواب خود
What is the effect of refractive index?	Affects (Refractive Index) = {?}	Optical Loss	دقیق	۱۰۰ %	> <b>KB:</b> Affects (Refractive Index) = {Optical Loss}
What are optical digital systems composed of?	Made Of (Optical Digital Systems) = {?}	Optical Components	قابل قبول	۸۵ %	> <b>ASPEC:</b> Optical Digital Systems > <b>KB:</b> Made Of (Optical Communication Systems) = {Optical Components}
How measure Material Dispersion?	Measured With (Material Dispersion) = {?}	Phase Shift Method	قابل قبول	۹۰ %	> <b>AGEN:</b> Intermodal Dispersion > <b>GEN:</b> Material Dispersion > <b>KB:</b> Measured With (Material Dispersion) = {Phase Shift Method}
How much is the photon loss of telecom C-band?	Photon Loss (Telecom C-band) = {?}	High	قابل قبول	۷۰ %	> <b>TDEP:</b> Absorption Loss +→ Photon Loss & N-H Bond Density +→ Absorption Loss → N-H Bond Density +→ Photon Loss > <b>KB:</b> N-H Bond Density(Telecom C-Band) = {High} > <b>DDEP:</b> N-H Bond Density(Telecom C-Band) = {High} → Photon Loss (Telecom C-Band) = {High}

در جدول (۲)، نوع جواب به پرسش اول "دقیق" انتخاب شده است، چراکه سامانه توانسته است آن را بطور مستقیم از درون پایگاه دانش و بدون انجام استدلال بدست آورد (عمق صفر). بنابراین درجه‌ی اطمینان حداکثر را برای آن منظور کرده است. برای جواب دوم، اگرچه سامانه تمام استنباطهای ممکن را آزمایش کرده اما تنها استنباط تخصیص آرگومان در عمق یک، نتیجه داده است. میزان اطمینان به این جواب بعد از ترکیب نتایج ۸۵ درصد محاسبه شده است. جواب سوم نیز دوباره در عمق استدلالی یک ولی با استنباط تعمیم آرگومان بدست آمده است. درجه اطمینان به این جواب از پاسخ دوم بیشتر بوده و علت آن این است که در الگوریتم ترکیب نتایج استدلال برای استنباطهای تعمیم، میزان اطمینان بیشتری در نظر گرفته شده

<sup>۱</sup> Backtrack

است. برای جواب آخر، در عمق دو با ترکیب استنباط های تعدی و اشتقاق از وابستگی بدست آمده است. سامانه ابتدا با استنباط تعدی اثبات کرده است که N-H Bond Density میتواند بصورت مستقیم روی Photon Loss تاثیر بگذارد و سپس یک استنباط اشتقاق از وابستگی نتیجه داده است که Photon Loss مربوط به محصول Telecom C-band زیاد است، از آنجاکه N-H Bond Density این محصول زیاد است.

کاربران خیره در حوزه مخابرات نیز بر قابل قبول بودن این جوابها صحه گذاشتند. مقبولیت این جوابها اعتبار الگوریتم پیاده سازی شده را تایید میکند لیکن باید آزمایشات بیشتر در مقیاس بزرگتر انجام شوند تا عملکرد الگوریتم استدلال در پایگاه های دانش غنی تر و با مجموعه پرسش های استاندارد نیز بررسی شود.

## ۹- نتیجه گیری و کارهای آینده

در این مقاله، یک مورد کاربردی از تئوری استدلال مقبول انسانی برای سامانه های جوابگویی پرسش ارائه شده و نتایج پیاده سازی آن در سامانه ی TeLQAS مورد بررسی قرار گرفته است. سامانه سعی میکند تا جواب دقیقی به پرسشهای مطرح شده بدهد و اگر این امر میسر نباشد، الگوریتم استدلال تلاش می کند تا جواب قابل قبولی با انجام زنجیره های استنباطی الهام گرفته شده از ذهن انسان بیابد.

انسانها دانش را از منابع گوناگون فرا می گیرند و این قابلیت ویژه را دارند تا آنها را براحتی با هم ترکیب کنند. در هنگام استدلال، فردی ممکن است دانش مربوط به حوزه، دانش عقل سلیم<sup>۱</sup> و دانش لغوی را بکار گیرد. اضافه کردن این دانش ها و پیدا نمودن نحوه صحیح ترکیب آنها می تواند قابلیت و کاربرد سامانه های پرسش و پاسخ را تا بسیار زیاد افزایش دهد. در ادامه-ی این تحقیق، سعی داریم که دانش لغوی را نیز به پایگاه دانش سامانه اضافه کنیم. یک راه انجام این کار در [۴] با معرفی روابط X و BN در منطق مقبول پیشنهاد شده است ولی مهمترین منبع برای این نوع دانش فرهنگهای دیجیتال<sup>۲</sup>، مانند WordNet [۱۰] هستند. ما تصمیم داریم در آینده هر دو روش را به خدمت بگیریم و در این راستا، ممکن است نیاز به توسعه ی تئوری استدلال مقبول باشد به نحوی که توصیفگرها نیز مانند آرگومانها و موارد ارجاع بتوانند تبدیل شوند.

## مراجع

- [1] A. Collins, R. Michalski, The logic of Plausible Reasoning A core theory. *Cognitive Science*, vol. 13, pp. 1-49, 1989.
- [2] K. Dontas, An Implementation of the Collins-Michalski Theory of Plausible Reasoning. Master's Thesis, University of Tennessee, Knoxville, TN, August 1987.
- [3] J.D. Kelly, PRS: A System for Plausible Reasoning, Master's Thesis, University of Illinois, Urbana-Champaign, 1989.
- [4] F. Oroumchian, R.N. Oddy, An Application of Plausible Reasoning to Information Retrieval. *SIGIR 1996*: 244-252.
- [5] F. Oroumchian, B.N. Araabi, E. Ashoori (2002). An Application of Plausible Reasoning and Dempster-Shafer Theory of Evidence in Information Retrieval. *In proceeding of International Conference on Fuzzy Systems and Knowledge Discovery*, pp. 442-447, 2002.
- [6] F. Oroumchian, B. Khandzah, Modeling an Intelligent Tutoring System by Plausible Inferences. *In proceeding of International Conference on Fuzzy Systems and Knowledge Discovery*, pp. 529-533, 2002.
- [7] A. Jalali, F. Oroumchian, An Evaluation of Document Clustering by means of Plausible Inferences, *International Journal of Computational Intelligence*, 2004.
- [8] M. Virvou, K. Kabassi, IFM: An Intelligent Graphical User Interface Offering Advice. *Preceding of 2<sup>nd</sup> Hellenic Conf. of AI, SETN 2002, Greece*, Companion Volume, pp. 155-164.
- [9] M.R. Hejazi, M. S. Mirian, K. Neshatian, A. Jalali, B.R. Ofoghi (2003). TeLQAS: A Telecommunication Literature Question Answering System Benefits from a Text Categorization Mechanism. *In proceeding of International Conference on Information and Knowledge Engineering*, USA, pp. 500-504, 2003.
- [10] C.D. Fellbaum (Ed.), *WordNet: An Electronic Lexical Database*. The MIT Press, Cambridge, London, England, 1998.

<sup>۱</sup> Commonsense knowledge

<sup>۲</sup> Digital Dictionary