دانشگاه سیستان وبلوچستان

تحصیلات تکمیلی

پایان نامه کارشناسی ارشد ریاضی کاربردی

گرایش تحقیق در عملیات

عنوان:

الگوریتم خوشه بندی کارا مبتنی بر مجموعه های فازی شهودی

استاد راهنما:

دکتر حسن میش مست نهی

تحقیق و نگارش:

مصطفی یگانه فر

بهمن ۱۳۹۰

چکیده

نظریه مجموعه های فازی شهودی به عنوان توسیعی از مجموعه های فازی دامنه کاربرد زیادی از قبیل برنامه ریزی های منطقی، تشخیص های پزشکی، الگوهای شناختی و موارد بسیاری از این دست را شامل می شود. استفاده از این مجموعه ها به عنوان ابزاری کارا و سودمند در زمینه تکنیک های خوشه بندی می تواند بسیار مفید باشد. همان طور که می دانیم الگوریتم های خوشه بندی گوناگونی برای استخراج دانش از درون اطلاعات مختلف وجود دارد.

یکی از کاربردی ترین این الگوریتم ها، الگوریتم $Fuzzy \ \ C-means$ است که به علت سادگی روش و حجم کم محاسبات، در چند سال گذشته بسیار مورد استفاده قرار گرفته است. اگرچه این روش با ضعف هایی نیز همراه است ولی با استفاده از الگوریتم های ترکیبی خوشه بندی تا حد زیادی می توان بر این ضعف ها غلبه کرد. ما در این پایان نامه با ترکیب الگوریتم خوشه بندی $Fuzzy \ \ C-means$ و نظریه مجموعه های فازی شهودی، الگوریتم خوشه بندی جدیدی را معرفی کرده و سپس کارایی این الگوریتم را با نتایج عددی نشان می دهیم. در انتها نیز پیشنهاداتی را جهت زمینه های کاری در آینده ارائه خواهیم کرد.

کلمات کلیدی: مجموعه های فازی شهودی، خوشه بندی فازی، الگوریتم C میانگین فازی، معیار تشابه شهودی

آزادی نه دادنی ست، نه گرفتنی ست، یادگرفتنی است

تقدیم به:

همسر مهربانم

آنکه وجودش همواره امیدی برای زندگیست و

سروش عزیزم

که با آمدنش زندگیمان را رنگ و بویی تازه بخشید .

تشکر و قدردانی :

خدایا

به من توفیق تلاش در شکست، صبر در ناامیدی، رفتن بی‌همراه، تلاش بی‌پاداش، دین بی‌دنیا، ایمان بی‌ریا، مناعت بی‌غرور و دوست داشتن بی آنکه دوست بداند روزی کن.

اینک خدا را به خاطر تمام نعمت‌ها و رحمت‌هایی که به بنده حقیر عطا کرده است سپاس می‌گویم و امیدوارم که قدردان الطاف بی‌کرانش باشم.

برخود لازم می‌بینم که از استاد راهنمای محترم جناب آقای دکتر حسن میش مست نهی که در تمام مراحل با صبر، شکیبایی و خلق نیکو مرا راهنمایی نموده‌اند تشکر و سپاسگذاری نمایم.

جا دارد از هم اتاقی های عزیزم آقایان مهندس افشین شاه حیدری، مهندس الیاس وحدت، مهندس میثم حسینی، مهندس بهرام طباطبایی، مهندس رضا جمشیدیان، مهندس شایان کویی نژاد، مهندس صادق صبوریان که خاطراتی خوبی را در این دانشگاه برایم رقم زدند، تشکر ویژه می نمایم.

همچنین از دوستان عزیزم آقایان محمد سیف پناه، بهنام بازیار، علی کریمی، عماد رجحانی، مهدی یادگاری و سرکار خانم عبدالهی که کمک هایشان را از این جانب دریغ ننمودند، تشکر می نمایم.

مصطفی یگانه فر

فهرست مندرجات

فهرست جداول

فهرست شکل‌ها

فصل ۱

مفاهیم مقدماتی

۱—۱ مقدمه

نظریه مجموعه های فازی برای اولین بار توسط پرفسور لطفی زاده[1] [۳۸] در سال ۱۹٦۵ با هدف مدل سازی ابهامات و عدم قطعیت رویدادها مطرح گردید. این نظریه عدم قطعیت احتمالی را پشتیبانی می کند و به دنبال ارائه قالبی برای نمایش و مدیریت دانش مبهم و عدم قطعی است. در نظریه مجموعه های فازی مرز قطعی و مشخصی برای مجموعه ها وجود ندارد و مانند مجموعه های دقیق نمی توان تابع مشخصه با برد صفر و یک تعریف نمود.

از آنجایی که منطق فازی با واقعیت ها و حقایقی که فکر بشر با آنها درگیر است، بیشتر منطبق است و از طرفی مدل سازی پدیده ها و فرآیندها با استفاده از نظریه فازی ساده تر و حقیقی تر به نظر می رسد، این نظریه در بسیاری از علوم و گرایش های گوناگون مورد استفاده قرار گرفته تا جایی که امروزه، گذشته از ابداع و تکمیل کامپیوترهایی مبتنی بر منطق فازی، این منطق در تجزیه و تحلیل داده ها و بسیاری از زمینه های دیگر نیز وارد شده است.

۱—۲ مفاهیم فازی

در علوم فنی و مهندسی و علوم پایه، دانشمندان برای مدل سازی پدیده ها یک رابطه به ظاهر دقیق ریاضی ارائه کرده اند، هر چند در کاربردهای عملی از دقت های آنچنان بالا استفاده نمی کنند ولی در دنیای پیرامون ما هر آنچه اتفاق می افتد، مدل ساده و ثابتی ندارد. در مسائلی که در عمرمان تجربه کرده ایم مواردی وجود دارد که هیچگاه نتوانسته و یا حتی نخواسته ایم معیار کمی و عددی روی آن بگذاریم. به عنوان مثال در محاورات روزمره از عباراتی مثل نزدیک یا دور برای بیان فاصله، گران یا ارزان برای ارزیابی قیمت استفاده می کنیم. اگرچه عبارت نزدیک اطلاعات دقیقی در خود ندارد ولی یک احساس ذهنی از میزان فاصله به شنونده ارائه کرده است و ذهن بر مبنای آن می تواند تصمیم گیری کند (این تصمیم گیری به هیچ عنوان بر

[1] Zadeh

مبنای احتمال نیست). در دنیای واقعی، کنترل و نظارت بسیاری از پدیده ها و فرآیندها به دقت نیاز دارند. حتی بسیاری از این پدیده ها و فرآیندها با مدل های ریاضی قابل بیان و توجیه نیستند. شاید نتوان عملکرد ماشینی را که با توجه به چند قاعده حسی و ذهنی بسیار ساده و با یک دقت کاملا معمولی توسط انسان کنترل می شود را با فرمول و قواعد ریاضی و محاسباتی مدل سازی و کنترل کرد.

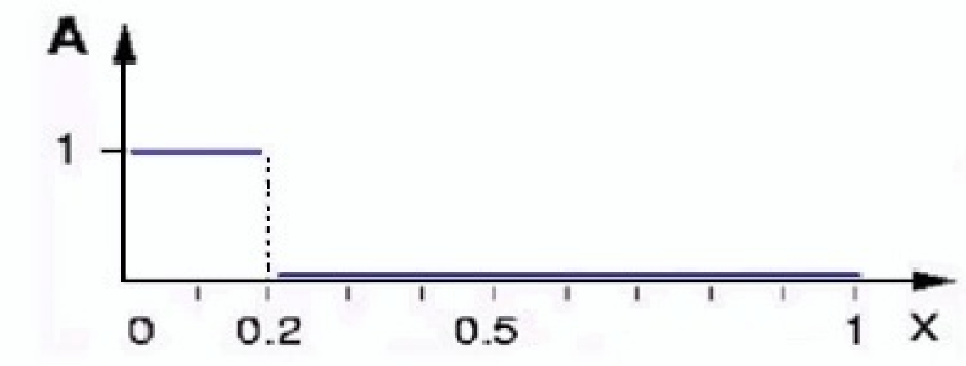
منطق فازی که توسط پروفسور لطفی زاده معرفی شده، مفهوم مجموعه ها، کمیت های نادقیق و همراه با عدم صراحت را در یک روال کاملا منظم و قانونمند معرفی کرده است و روش های تصمیم گیری در چنین محیط هایی را پیشنهاد می کند. از آنجایی که منطق فازی با واقعیت ها و حقایقی که فکر بشر با آنها درگیر است، بیشتر منطبق است و از طرفی مدل سازی پدیده ها و فرآیند ها با استفاده از نظریه فازی ساده تر و حقیقی تر به نظر می رسد. در نظریه مجموعه های فازی، اصول و قوانینی تعریف شده است که مجموعه های کلاسیک را نیز به عنوان یک حالت خاص از مجموعه های فازی در بر می گیرد. در مجموعه های کلاسیک تابع عضویت فقط مقادیر صفر و یک را شامل می شود (یک شی یا عضو یک مجموعه هست یا نیست) ولی در مجموعه های فازی این تابع به صورت درجه عضویت بیان می شود (هر عضو در یک مجموعه فازی می تواند هم عضو آن مجموعه باشد و هم نباشد).

اصولا یک مجموعه فازی به صورت مجموعه ای از عناصر در جهان اطلاعات تعریف می شود به صورتی که حدود مجموعه ای که در این مجموعه جهانی قرار دارد مبهم یا به عبارتی فازی است. در مجموعه کلاسیک، تابع عضویت تابعی است به اصطلاح تیز و مشخص با دو مقدار صفر و یک. به عبارت واضحتر یک عضو x یا به مجموعه ای مثل $\mathbf{A}$ تعلق دارد یا ندارد و از این دو حالت خارج نیست. ولی برای یک مجموعه فازی تابع عضویت تابعی است با تغییر تدریجی و نه آنی. یعنی مفهوم یا هست و یا نیست کنار گذاشته و از مفهوم میزان و درجه عضویت (چقدر هست و چقدر نیست) استفاده می شود. در حقیقت محدوده یک مجموعه فازی مبهم و مشکوک است و در یک کلام فازی است. برای روشنتر شدن موضوع فرض کنید بخواهیم بین اعضای یک جامعه از نظر سنی، مجموعه افراد نوجوان و جوان و مسن را تعریف کنیم. محدوده بین نوجوانی و جوانی و پیری را نمی توان به صورت دقیق مشخص کرد و هیچ محدوده صریح و مطمئنی برای ارزیابی وجود ندارد. نظریه مجموعه های فازی از همین مفهوم نشات می گیرد. گذشته از معیارهایی که اصلا نمی

توان آنها را اندازه گیری کرد (مانند زیبایی و زشتی، خوبی و بدی) بسیاری از کمیت ها در زندگی روزمره موجودند که برای ما مقدار آنها اهمیتی ندارد و برای بیان آنها به طور ناخود آگاه از عبارات فازی استفاده می کنیم (مانند سردی و گرمی هوا، نزدیکی یا دوری مقصد). جالب اینجاست که همین عبارات فازی که هیچ گونه اطلاعات کمی دقیق در خود ندارند به نحو موثری در تصمیم گیری های روزمره بکار می آیند.

۱—۳ مدل فازی متغیرها

در ریاضیات کلاسیک با مجموعه های قطعی (غیر فازی) آشنا شده ایم. برای مثال فرض کنید که **U** مجموعه اعداد حقیقی بین صفر و یک باشد. می توان یک زیر مجموعه از **U** به نام **A** بدین صورت تعریف کرد: مجموعه مقادیر کوچکتر و یا مساوی ۰/۲. در این صورت تابع مشخصه **A** در شکل (۱—۱) نشان داده شده است.



شکل (۱—۱): تابع مشخصه مجموعه غیر فازی **A**.

مقدار این تابع برای مقادیری از x که عضو **A** باشند یک و برای بقیه مقادیر صفر می باشند. مقادیری که مقدار تابع را یک می کنند میتوان به صورت مقادیری که عضو **A** هستند و مقادیری که تابع را صفر می کنند، مقادیری که عضو **A** نیستند بیان نمود. بنابراین میتوان گفت ۰/۲ عضو این مجموعه است اما ۰/۷ عضو این مجموعه نیست. اما همانطور که مشخص است این تابع انعطاف پذیری کمی دارد برای مثال اگر بخواهیم

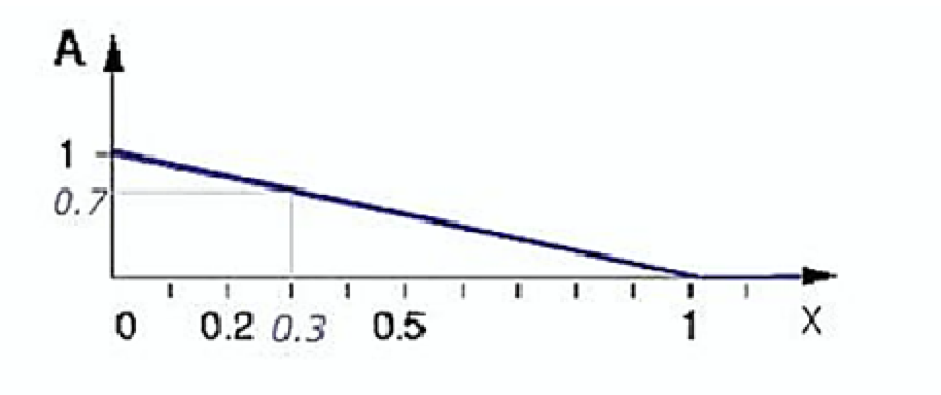
اعداد نزدیک به صفر را نمایش بدهیم با مشکل مواجه می شویم. یک جنبه این است که نمی توانیم اعضای مجموعه را بیان کنیم و جنبه دیگر اینکه مرز مشخصی برای عضویت و یا عدم عضویت در این مجموعه وجود ندارد. برای حل این مشکل از مجموعه های فازی کمک می گیریم. در ابتدا تعریفی از مجموعه های فازی را ارائه خواهیم کرد:

تعریف ۱—۱: یک مجموعه فازی $\tilde{A}$ از مجموعه مرجع $\mathbf{X}$ اینگونه تعریف می شود:

$$\tilde{A} = \{\langle x, \mu_{\tilde{A}}(x)\rangle \mid \ x \in X\}$$

که $\mu_{\tilde{A}} : \mathbf{X} \to \ [\mathbf{0}, \mathbf{1}]$ تابع عضویت مجموعه فازی $\tilde{A}$ است و $\mu_{\tilde{A}}(x) \in [۰, ۱]$ درجه عضویت $x \in \mathbf{X}$ در $\tilde{A}$ است.

منطق فازی اجازه می دهد درجه عضویت هر عنصر عددی بین صفر و یک ([۰, ۱]) باشد. در این حالت تابع مشخصه ای به نام تابع تعلق داریم که میتواند هر مقداری در بازه [۰, ۱] را اختیار کند. بنابراین میتوان تابع تعلق ($\mu(x) = ۱ - x, \quad ۰ \leq x \leq ۱$) را برای زیرمجموعه یاد شده آورد و همانطور که در شکل (۱—۲) دیده می شود در این حالت می توان گفت که عدد ۰/۳ با درجه عضویت ۰/۷ متعلق به مجموعه اعداد نزدیک به صفر می باشد.



شکل (۱—۲): مجموعه فازی اعداد نزدیک به صفر.

خواص و ویژگی هایی که برای تعیین اعضای مجموعه فازی بیان می شوند به صورت فازی هستند و یک توصیف دقیق نمی باشند، بنابراین می توان از توابع تعلق مختلف برای نشان دادن یک مجموعه فازی استفاده

کرد. در عمل منحنی هایی به کار می رود که نمایش ریاضی ساده ای داشته باشند و با تعداد پارامتر کمی قابل تنظیم باشند.

۱—۴ عملیات بر روی مجموعه های فازی

همانند مجموعه های قطعی، برای مجموعه های فازی نیز عملگر های مکمل، اجتماع و اشتراک وجود دارد که در ادامه به بررسی آنها می پردازیم.

۱—۴—۱ عملگر مکمل

تابع تعلق مکمل مجموعه فازی **A** را بصورت $\mu_{\bar{A}}(x)$ نشان می دهیم. تابع مکمل باید بتواند چند شرط زیر را برآورده سازد:

$$\mu_A(x) = ۰, \Rightarrow \mu_{\bar{A}}(x) = ۱$$
$$\mu_A(x) = ۱, \Rightarrow \mu_{\bar{A}}(x) = ۰$$
$$\mu_A(x) < \mu_A(y) \Rightarrow \mu_{\bar{A}}(x) \geq \mu_{\bar{A}}(y)$$

توابع زیادی شروط فوق را برآورده می سازند که یکی از معروفترین آنها $\mu_{\bar{A}}(x) = ۱ - \mu_A(x)$ می باشد.

۱—۴—۲ عملگر اجتماع

عملگر اجتماع را با **S** نشان داده و بصورت زیر نمایش می دهند:

$$S[\mu_A(x), \mu_B(x)] = \mu_{A \cup B}(x)$$

به این معنی که $S : [۰, ۱] \times [۰, ۱] \to [۰, ۱]$ نگاشتی است که توابع تعلق **A** و **B** را به تابع تعلق اجتماع **A** و **B** تبدیل می کند. عملگر اجتماع باید شرط های زیر را برآورده سازد:

۱) شرط های مرزی زیر در آن صدق کند :

$$S[\,0,\,\mu_A(x)] = S[\,\mu_A(x),\,0] = \mu_A(x)$$
$$S[1,1] = 1$$

۲) دارای شرط جابه جایی باشند:

$$S[\,\mu_A(x),\,\mu_B(x)] = S[\,\mu_B(x),\,\mu_A(x)]$$

۳) شرط یکنوایی را دارا باشد:

$$\left.\begin{array}{l}\mu_{A_1}(x) < \mu_{A_2}(x)\\ \mu_{B_1}(x) < \mu_{B_2}(x)\end{array}\right\}\mu_{A_1\cup B_1}(x) < \mu_{A_2\cup B_2}(x)$$

۴) شرط شرکت پذیری در آن صدق کند:

$$\mu_{(A\cup B)\cup C}(x) = \mu_{A\cup(B\cup C)}(x)$$

۱–۴–۳ عملگر اشتراک

عملگر اشتراک را با t نشان داده و بصورت زیر نمایش می دهند.

$$t\,[\,\mu_A(x),\,\mu_B(x)] = \mu_{A\cap B}(x)$$

یعنی $t\ :\ [0,1]\times[0,1]\rightarrow[0,1]$ تابعی است که توابع تعلق مجموعه های فازی **A** و **B** را به تابع تعلق اشتراک **A** و **B** تبدیل می کند. برای اینکه رابط t عملگر اشتراک باشد باید چهار شرط زیر را برآورده نماید:

۱) شرط مرزی

$$t\,[0,0] = 0,\quad t[\mu_A(x),1] = t[1,\mu_A(x)] = \mu_A(x)$$

۲) شرط جابجایی

$$t[\mu_A(x),\,\mu_B(x)] = t[\,\mu_B(x),\mu_A(x)]$$

۳) شرط یکنوایی

$$\left.\begin{array}{l}\mu_{A_1}(x) < \mu_{A_2}(x)\\ \mu_{B_1}(x) < \mu_{B_2}(x)\end{array}\right\} \quad \mu_{A_1\cap B_1}(x) < \mu_{A_2\cap B_2}(x)$$

۴) شرط شرکت پذیری

$$\mu_{(A\cap B)\cap C}(x) = \mu_{A\cap(B\cap C)}(x)$$

پیشنهاد پروفسور زاده، عملگر min برای اشتراک بوده است که به صورت زیر بیان می شود:

$$\mu_{A\cap B}(x) = \min[\mu_A(x), \mu_B(x)]$$

بنابر کاربردهای مختلف، میتوان عملگرهای اجتماع و اشتراک مختلفی تعریف کرد که مطابق آنچه گفته شد باید شرایط ذکر شده را برآورده سازد. برای مشاهده روابط و تعاریف انواع مختلف عملگرهای فوق به مراجع [۱۸و۳۵] مراجعه کنید.

۱ــ۵ مجموعه های فازی شهودی

منطق و مجموعه های فازی شهودی برای نخستین بار توسط آتاناسوف[1] [۲] معرفی شد. پس از آن مقالات زیادی در مفهوم منطق فازی شهودی و اعمال روی آن ها توسط شخص آتاناسوف و سایرین نوشته شده است[۲و۴۱]. در این بخش ما به اختصار راجع به مجموعه های فازی شهودی بحث خواهیم کرد و اعمال روی این مجموعه ها را تعریف می کنیم. حال در ادامه تعریفی تعمیم یافته از مجموعه های فازی را که توسط آتاناسوف ارائه شده است را در زیر بیان می کنیم.

تعریف ۱ــ۲: یک مجموعه فازی شهودی **A** از مجموعه مرجع **X** اینگونه تعریف می گردد:

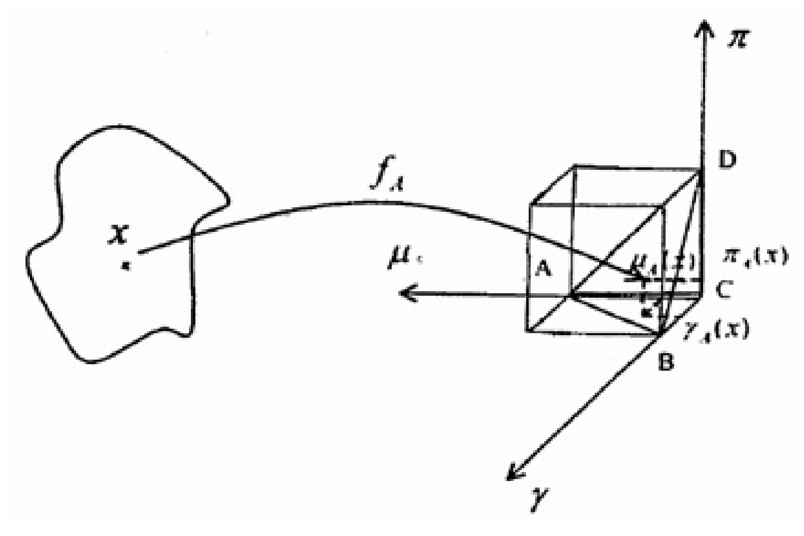
$$A = \{\langle x, \mu_A(x), \gamma_A(x)\rangle \mid \ x \in X\}$$

[1] K.Attanassov

که $\mu_A : X \to [۰,۱]$ و $\gamma_A : X \to [۰,۱]$ با این شرط که $۰ \le \mu_A(x) + \nu_A(x) \le ۱$ که مقادیر حقیقی $\mu_A(x)$ و $\gamma_A(x)$ که متعلق به بازه $[۰,۱]$ می باشند، درجه عضویت و درجه عدم عضویت x به $\mathbf{A}$ نامیده می شود.

هر مجموعه فازی $\tilde{\mathbf{A}}$ حالت خاصی از مجموعه های فازی شهودی است و می توان آن را به شکل مجموعه فازی شهودی $A = \{\langle x, \mu_{\tilde{A}}(x), ۱ - \mu_{\tilde{A}}(x) \rangle \mid x \in X\}$ نشان داد. برای هر مجموعه فازی شهودی $\mathbf{A}$ از $\mathbf{X}$، $\pi_A(x) = ۱ - \mu_A(x) - \gamma_A(x)$ را درجه تردید یا شاخص شهودی x در $\mathbf{A}$ می نامیم. این در حقیقت درجه تعریف x در $\mathbf{A}$ است. بوضوح برای هر x متعلق به $\mathbf{A}$ داریم $۰ \le \pi_A(x) \le ۱$.

برای هر مجموعه فازی $\tilde{\mathbf{A}}$ از $\mathbf{X}$ داریم $\forall x \in X : \quad \pi_{\tilde{A}}(x) = ۱ - \mu_{\tilde{A}}(x) - [۱ - \mu_{\tilde{A}}(x)] = ۰$ یک شمای هندسی از مجموعه های فازی شهودی اینگونه است:



شکل (۱—۳): تفسیر هندسی از یک مجموعه فازی شهودی $\mathbf{A}$ از $\mathbf{X}$.

در شکل (۱—۳) مجموعه فازی شهودی $\mathbf{A}$ از $\mathbf{X}$ به کمک تابع f_A که به هر دو مقدار $\mu_A(x)$ و $\gamma_A(x)$ نسبت می دهد، مشخص می شود. تصویر نقطه $f_A(x)$ از مثلث $\mathbf{ABD}$ روی محور $\mathbf{AC}$ درجه عضویت x در $\mathbf{A}$، $\mu_A(x)$ و تصویر آن بر روی محور $\mathbf{BC}$ درجه عدم عضویت x در A، $\gamma_A(x)$ و تصویر آن روی محور $\mathbf{DC}$ درجه تردید x در $\mathbf{A}$، $\pi_A(x) = ۱ - \mu_A(x) - \gamma_A(x)$ است. بنابراین به هر عضو از مجموعه فازی شهودی،