



جمهوری اسلامی ایران
Islamic Republic of Iran

سازمان ملی استاندارد ایران

Iranian National Standardization Organization



استاندارد ملی ایران

۱۶۷۶۹

چاپ اول

مهر ۱۳۹۲

INSO

16769

1st. Edition

Oct.2013

مدیریت منابع زبان - چارچوب
حاشیه‌نویسی صرفی (SynAF)

Language resource management-
Syntactic annotation framework (SynAF)

ICS:01.020

به نام خدا

آشنایی با سازمان ملی استاندارد ایران

مؤسسه استاندارد و تحقیقات صنعتی ایران به موجب بند یک ماده ۳ قانون اصلاح قوانین و مقررات مؤسسه استاندارد و تحقیقات صنعتی ایران، مصوب بهمن ماه ۱۳۷۱ تنها مرجع رسمی کشور است که وظیفه تعیین، تدوین و نشر استانداردهای ملی (رسمی) ایران را به عهده دارد. نام مؤسسه استاندارد و تحقیقات صنعتی ایران به موجب یکصد و پنجاه و دومین جلسه شورای عالی اداری مورخ ۹۰/۶/۲۹ به سازمان ملی استاندارد ایران تغییر و طی نامه شماره ۲۰۶/۳۵۸۳۸ مورخ ۹۰/۷/۲۴ جهت اجرا ابلاغ شده است.

تدوین استاندارد در حوزه های مختلف در کمیسیون های فنی مرکب از کارشناسان سازمان، صاحب نظران مراکز و مؤسسات علمی، پژوهشی، تولیدی و اقتصادی آگاه و مرتبط انجام می شود و کوششی همگام با مصالح ملی و با توجه به شرایط تولیدی، فناوری و تجاری است که از مشارکت آگاهانه و منصفانه صاحبان حق و نفع، شامل تولیدکنندگان، مصرف کنندگان، صادرکنندگان و وارد کنندگان، مراکز علمی و تخصصی، نهادها، سازمان های دولتی و غیر دولتی حاصل می شود. پیش نویس استانداردهای ملی ایران برای نظرخواهی به مراجع ذی نفع و اعضای کمیسیون های فنی مربوط ارسال می شود و پس از دریافت نظرها و پیشنهادهای در کمیته ملی مرتبط با آن رشته طرح و در صورت تصویب به عنوان استاندارد ملی (رسمی) ایران چاپ و منتشر می شود.

پیش نویس استانداردهایی که مؤسسات و سازمان های علاقه مند و ذی صلاح نیز با رعایت ضوابط تعیین شده تهیه می کنند در کمیته ملی طرح و بررسی و در صورت تصویب، به عنوان استاندارد ملی ایران چاپ و منتشر می شود. بدین ترتیب، استانداردهایی ملی تلقی می شوند که بر اساس مفاد نوشته شده در استاندارد ملی ایران شماره ۵ تدوین و در کمیته ملی استاندارد مربوط که سازمان ملی استاندارد ایران تشکیل می- دهد به تصویب رسیده باشد.

سازمان ملی استاندارد ایران از اعضای اصلی سازمان بین المللی استاندارد (ISO)^۱، کمیسیون بین المللی الکتروتکنیک (IEC)^۲ و سازمان بین المللی اندازه شناسی قانونی (OIML)^۳ است و به عنوان تنها رابط^۴ کمیسیون کدکس غذایی (CAC)^۵ در کشور فعالیت می کند. در تدوین استانداردهای ملی ایران ضمن توجه به شرایط کلی و نیازمندی های خاص کشور، از آخرین پیشرفت های علمی، فنی و صنعتی جهان و استانداردهای بین المللی بهره گیری می شود.

سازمان ملی استاندارد ایران می تواند با رعایت موازین پیش بینی شده در قانون، برای حمایت از مصرف کنندگان، حفظ سلامت و ایمنی فردی و عمومی، حصول اطمینان از کیفیت محصولات و ملاحظات زیست محیطی و اقتصادی، اجرای بعضی از استانداردهای ملی ایران را برای محصولات تولیدی داخل کشور و/یا اقلام وارداتی، با تصویب شورای عالی استاندارد، اجباری نماید. سازمان می تواند به منظور حفظ بازارهای بین المللی برای محصولات کشور، اجرای استانداردهای کالاهای صادراتی و درجه بندی آن را اجباری نماید. همچنین برای اطمینان بخشیدن به استفاده کنندگان از خدمات سازمان ها و مؤسسات فعال در زمینه مشاوره، آموزش، بازرسی، ممیزی و صدور گواهی سیستم های مدیریت کیفیت و مدیریت زیست محیطی، آزمایشگاه ها و مراکز کالیبراسیون (واسنجی) وسایل سنجش، سازمان ملی استاندارد ایران این گونه سازمان ها و مؤسسات را بر اساس ضوابط نظام تأیید صلاحیت ایران ارزیابی می کند و در صورت احراز شرایط لازم، گواهینامه تأیید صلاحیت به آن ها اعطا و بر عملکرد آن ها نظارت می کند. ترویج دستگاه بین المللی یکاها، کالیبراسیون (واسنجی) وسایل سنجش، تعیین عیار فلزات گرانبها و انجام تحقیقات کاربردی برای ارتقای سطح استانداردهای ملی ایران از دیگر وظایف این سازمان است.

1- International Organization for Standardization

2 - International Electrotechnical Commission

3- International Organization of Legal Metrology (Organisation Internationale de Metrologie Legale)

4 - Contact point

5 - Codex Alimentarius Commission

فهرست مندرجات

صفحه	عنوان
ب	آشنایی با سازمان ملی استاندارد
ج	کمیسیون فنی تدوین استاندارد
د	پیش گفتار.....
هـ	مقدمه.....
۱	هدف و دامنه کاربرد
۱	۲ مراجع الزامی.....
۱	۳ اصطلاحات و تعاریف
۸	۴ فرآمدل SynAF
۸	۱-۴ کلیات
۹	۲-۴ فرآمدل SynAF
۹	۱-۲-۴ رئوس کلی.....
۱۱	۲-۲-۴ رده گره صرفی.....
۱۱	۳-۲-۴ رده <i>T-Node</i>
۱۱	۴-۲-۴ رده <i>NT-Node</i>
۱۱	۵-۲-۴ رده لبه نحوی
۱۲	۶-۲-۴ رده حاشیه نویسی.....
۱۳	پیوست الف (الزامی) رده داده‌ها برای SynAF.....
۳۰	پیوست ب (اطلاعاتی) ارتباط چارچوب حاشیه نویسی زبانی.....
۳۳	پیوست پ (اطلاعاتی) کتابنامه

کمیسیون فنی تدوین استاندارد

« مدیریت منابع زبان - چارچوب حاشیه نویسی صرفی (SynAF) »

رئیس:

احمدی، عباس

(کارشناس ارشد زبان شناسی)

دبیر:

علی بیگی، محمود

(کارشناسی مترجمی زبان انگلیسی)

اعضاء: (به ترتیب حروف الفبا)

ابراهیمی، محمد

(کارشناسی مترجمی زبان انگلیسی)

بسطامی، مهدی

(کارشناسی مترجمی زبان انگلیسی)

خانی، بهزاد

(کارشناس زبان و ادبیات فارسی)

مظلوم، فاطمه

(کارشناسی میکروبیوشمی)

نعمتی، معصومه

(کارشناس زبان انگلیسی)

سمت و/یا نمایندگی

مدرس دانشگاه آزاد واحد ایلام

مدیر مرکز آموزش‌های کوتاه مدت جهاد

دانشگاهی واحد استان ایلام

کارشناس بانک صادرات شعبه سعدی

کارشناس بانک سپه شعبه فجر

کارشناس جهاددانشگاهی واحد استان ایلام

کارشناس اداره کل استاندارد و تحقیقات

صنعتی استان ایلام

روابط عمومی اداره کل استاندارد و تحقیقات

صنعتی استان ایلام

پیش‌گفتار

استاندارد "مدیریت منابع زبان - چارچوب حاشیه‌نویسی صرفی (SynAF)" که پیش‌نویس آن در کمیسیون‌های مربوط تهیه و تدوین شده و در هشتاد و نهمین اجلاس کمیته ملی استاندارد اسناد و تجهیزات اداری و آموزشی مورخ ۸/۱۱/۱۳۹۰ مورد تصویب قرار گرفته است، اینک به استناد بند یک ماده ۳ قانون اصلاح قوانین و مقررات مؤسسه استاندارد و تحقیقات صنعتی ایران، مصوب بهمن ماه ۱۳۷۱، به عنوان استاندارد ملی ایران منتشر می‌شود.

برای حفظ همگامی و هماهنگی با تحولات و پیشرفت‌های ملی و جهانی در زمینه صنایع، علوم و خدمات، استانداردهای ملی ایران در مواقع لزوم تجدید نظر خواهد شد و هر پیشنهادی که برای اصلاح و تکمیل این استانداردها ارائه شود، هنگام تجدید نظر در کمیسیون فنی مربوط مورد توجه قرار خواهد گرفت. بنابراین، باید همواره از آخرین تجدید نظر استانداردهای ملی استفاده کرد.

منبع و مآخذی که برای تهیه این استاندارد مورد استفاده قرار گرفته به شرح زیر است:

ISO 24615: 2010, Language resource management - Syntactic annotation Framework (SynAF)

مقدمه

این استاندارد ملی ایران بر اساس پروژه‌های متعدد و فعالیت‌هایی پیش از استاندارد سازی است، که در چند سال گذشته (به [9] Abeillé 2001 مراجعه کنید) صورت گرفته است برای ارائه مدل‌ها و قالب‌های مرجع جهت نشان دادن اطلاعات صرفی، چه به عنوان خروجی تجزیه‌گر صرفی، چه به عنوان حاشیه نویسی منابع زبان (بانک‌های درختی). برای چندین سال ابتکار بانک درختی پن (Penn Treebank) به عنوان یک استاندارد بالفعل برای بانک داری درختی مورد استفاده قرار گرفته است، ولیکن کارهای اخیر برای مثال: ابتکار /Negra در آلمان یا ابتکار ISST در ایتالیا (به پیوست کتابشناسی بند ۱۸ Montemagni مراجعه کنید) تداوم چارچوب منسجم تری را نشان می‌دهد، که می‌تواند پاسخ‌گویی باشد برای هر دو پدیده سازواری (سلسله مراتبی) و وابستگی در حاشیه نویسی صرفی.

پروژه eContent "LIRICS"، نقطه عطفی بوده است در گردآوری گروهی از کارشناسانی که تدوین استاندارد ISO 24615 (SynAF) را به عهده داشتند. این گروه در حالی که SynAF را آماده می‌کردند، تایید کردند که طرح‌های ابتکاری موجود در واقع یک مدل داده‌های عمومی را به اشتراک می‌گذارند که اساس خوبی برای فرآمدل SynAF ارائه می‌دهد (به مطالعه ساخته شده در خروجی D.3.1 «ارزیابی طرح‌های ابتکاری برای حاشیه نویسی صرفی - نحوی و صرفی» از پروژه اروپایی LIRICS مراجعه کنید).

این استاندارد ملی ایران فرآمدلی برای حاشیه نویسی صرفی به همراه یک فهرست از رده‌اطلاعات مربوط به حاشیه نویسی صرفی را ارائه می‌دهد. مقوله‌های داده‌ای بر روی خدمات رسان ISOCat (<http://www.isocat.org>) در مشخصات نحو (همان طور که در ایزو ۲۰۰۹:۱۲۶۲۰ تعریف می‌شوند) در دسترس می‌باشند.

ISO 24615: 2010, Language resource management — Syntactic annotation Framework (SynAF)

مدیریت منابع زبان - چارچوب حاشیه‌نویسی صرفی (SynAF)

۱ هدف و دامنه کاربرد

هدف از تدوین این استاندارد توصیف چارچوب حاشیه‌نویسی صرفی (SynAF)، که مدلی پیشرفته است برای نشان دادن حاشیه‌نویسی صرفی داده‌های زبانی، به منظور ایجاد مشارکت میان منابع زبان یا اجزای پردازش زبان.

این استاندارد ملی در ارتباط نزدیک و مکمل استاندارد ISO ۲۴۶۱۱ (MAF)، چارچوب حاشیه‌نویسی صرفی - نحوی) است و فرآمدلی را برای نمادهای صرفی و هم‌چنین رده‌های داده‌های مرجع برای نشان دادن اطلاعات تشکیل دهنده و وابسته و در جملات یا دیگر سخنان و قطعات قابل قیاس ارائه می‌دهد.

۲ مراجع الزامی

مدارک الزامی زیر حاوی مقرراتی است که در متن این استاندارد ملی ایران به آن‌ها ارجاع داده شده است. بدین ترتیب آن مقررات جزء این استاندارد ملی ایران محسوب می‌شود. در صورتی که به مدرکی با ذکر تاریخ انتشار ارجاع داده شده باشد، اصلاحیه‌ها و تجدید نظرهای بعدی آن مورد نظر این استاندارد ملی ایران نیست. در مورد مدارکی که بدون ذکر تاریخ انتشار به آن‌ها ارجاع داده شده است، همواره آخرین تجدیدنظر و اصلاحیه‌های بعدی آن‌ها مورد نظر است. استفاده از مراجع زیر برای این استاندارد الزامی است:

- 2.1 ISO 1087-1:2000, Terminology work - Vocabulary - Part 1: Theory and application
- 2.2 ISO 1087-2:2000, Terminology work - Vocabulary - Part 2: Computer applications
- 2.3 ISO 12620:2009, Terminology and other language and content resources - Specification of data categories and management of a Data Category Registry for language resources
- 2.4 ISO 24611, Language resource management — Morpho-syntactic annotation framework

۳ اصطلاحات و تعاریف

در این استاندارد علاوه بر اصطلاحات و تعاریف تعیین شده در ISO 1087-1, ISO 1087-2, ISO 12620، ISO 24611 اصطلاحات و تعاریف زیر نیز به کار می‌رود.

۱-۳

ادات^۱

عنصری در ساخت یک بند که جزئی از هسته آن بند قلمداد نمی‌شود (۱۹-۳).

یادآوری - قیدها ضمایم احتمالی برای یک جمله هستند.

۱

۲-۳

قطعه^۲

سازه (جزء تشکیل دهنده) غیر بازگشتی (۴-۳) است.

۳-۳

بند

گروهی از عبارات است (۱۴-۳) که معمولاً شامل یک گزاره است .

یادآوری - یک بند می‌تواند یک بند اصلی (۱۰-۳) یا بند فرعی (۱۷-۳) باشد. در زبان‌هایی که در مورد صرف کردن افعال تمایز

قائل هستند، بندهائی که گزاره آن‌ها یک فعل است، بسته به شکل فعل می‌توانند قابل صرف کردن یا غیرقابل صرف کردن باشند.

بند اصلی به تنهایی می‌تواند یک جمله کامل (۱۵-۳) بسازد. در مدل SynAF، یک بند حالت خاصی از یک سازه است (۴-۳).

۴-۳

سازه^۳

هر یک از اجزای تشکیل دهنده جمله در سطوح مختلف [به عبارات (۱۴-۳)]، عبارات [به بند (۳-۳)] یا عبارات دیگر [یا بندها [به جمله (۱۵-۳)]] بر اساس خواص ساختاری (یا سلسله مراتبی) است.

۵-۳

وابستگی^۴

رابطه وابستگی

1- Adjunct

2-Chunk از واحدهای نحوی یا معنایی زبان است که در عین حال ساختمان درونی خاص خود را دارد

3- Constituent

4- Dependency

رابطه صرفی میان صورت های واژگانی (۳-۲۴) یا سازه‌ها (۳-۴) بر اساس نقش دستوری (۳-۷) است، که سازه‌ها در ارتباط با یک دیگر ایفا می کنند

۳-۶

لبه نحوی^۱

لبه

سه گانه‌ای است با یک گره^۲ مبداء (۳-۱۲)، یک گره هدف، و حاشیه نویسی‌های اختیاری (۳-۹).

یادآوری - گره‌های غیر پایانی^۳ (۳-۱۳) دارای یک سازواری^۴ رونده لبه صرفی هستند.

۳-۷

نقش دستوری

یک صورت واژگانی (۳-۲۴) یا سازه (۳-۴) در محیط صرفی درونه‌ای آن است.

یادآوری - برای مثال: یک عبارت اسمی (NP) می تواند به عنوان یک فاعل در یک جمله عمل کند (۳-۱۵)، یا اسم ممکن است به عنوان وابسته فاعل فعل در نمودار وابستگی عمل کند. رابطه ای دستوری بین فاعل - عبارت اسمی - و فعل اصلی در جمله وجود دارد. همه روابط دستوری (موضوع - گزاره، هسته - معرف، و غیره) تحت مفهوم روابط وابستگی قرار می گیرند (۳-۵)، چه بین گره‌های پایانی یا غیر پایانی باشند.

۳-۸

هسته صرفی^۴

هسته

بخشی از یک سازه (۳-۴) است، که توزیع (محیط صرفی که ممکن است سازه در آن نمایان شود) و خواص دستوری آن را (برای مثال: اگر جنس دستوری هسته مونث باشد، پس جنس تمام سازه مونث خواهد شد) تعیین می کند.

1- Syntactic edge

2- Node

3 -Non-terminal node

4- Syntactic head

یادآوری - هسته سازه معمولاً نمی تواند حذف شود.

۹-۳

حاشیه نویسی زبان شناسی^۱

حاشیه نویسی

جفت مشخصه- ارزش^۲ خاصیتی زبانی به یک جزء زبانی می دهد.

۱۰-۳

بند اصلی^۳

بندی(۳-۳) است که می تواند به خودی خود خ^۳ ه عنوان یک جمله کامل (۳-۱۵) عمل کند.

یادآوری - در زبان هائی که محدودیت را در نظر می گیرند، بند اصلی معمولاً محدود است برای مثال: آن قطار دیر کرده است.

۱۱-۳

معرف^۴

بخشی از سازه است (۳-۴)، که خاصیتی را به هسته سازه نسبت می دهد (۳-۸).

یادآوری - معرف می تواند پیش یا پس از هسته عبارت قرار گیرد(۳-۱۴) (پیش معرف^۵ یا پس معرف^۶). معرف ها در سازه اختیاری می باشند.

۱۲-۳

گره^۷

گره صرفی^۸

1- Linguistic annotation

2- Feature-value pair

3- Main clause

4- Modifier

5- Pre-modifier

6-Post-modifier

7-Node

8- Syntactic node

صورت واژگانی (۲۴-۳) یا سازه (۴-۳)، که به عنوان جزء ابتدایی صرفی یک تجزیه و تحلیل صرفی به شمار می‌آید.

۱۳-۳

گره غیر پایانی

گره صرفی (۱۲-۳) است که صورتی واژگانی (۲۴-۳) نیست.

یادآوری - یک گره غیر پایانی دارای لبه سازواری صادرشونده (۶-۳) است .

۱۴-۳

عبارت

گروهی از صورت‌های واژگانی (۲۴-۳) معمولاً حاوی یک یا چند کلمه) است، که می‌تواند جای یک نقش دستوری را پر کند (۷-۳)، برای مثال در بند (۳-۳) .

یادآوری - عبارات خالی مجاز می‌باشند (ضمایر غیر ادراکی بودن که گاهی اوقات به عنوان "ضمیر" مشخص شده و داشتن نقش فاعل در بند). عبارت معمولاً با نام هسته آن نام گذاری می‌شود (۸-۳)، برای مثال: عبارت اسمی، عبارت فعلی، عبارت صفتی، عبارت قیدی و عبارات حرف اضافه نام گرفته است. عبارات به صورت غیر رسمی به عنوان "واژه‌های پف کرده" شناخته می‌شوند، که در آن بخش‌هایی از عبارت که به هسته اضافه شده اند، مرجع هسته را مشخص نموده و شرح می‌دهند. در مدل ما عبارت، مورد خاصی از سازه (۴-۳) است.

۱۵-۳

جمله

گروه مرتبگی از صورت‌های واژگانی (۲۴-۳) است، که حاوی مسند بوده، معمولاً یک اندیشه کامل را بیان می‌نماید و واحد اصلی ساختار گفتار را تشکیل می‌دهد.

یادآوری - جمله شامل یک یا چند بند (۳-۳) است. وقتی که گفتار توصیف می‌شود، معمولاً در مورد "گفته" به جای جملات صحبت می‌شود.

۱۶-۳

گستره^۱

جفتی از نقاط (p_2, p_1) ، که $p_2 \geq p_1$ است، بخشی از سند که حاشیه نویسی (۳-۹) بر روی آن اعمال می شود را شناسایی می کند.

یادآوری - گستره چندگانه، رشته ای از گستره ها است که در آن نقطه پایانی هر گستره کم تر یا مساوی با نقطه آغازین گستره بعدی است.

۱۷-۳

بند پیرو^۲

بندی است، که نقش دستوری (۳-۷) را در یک عبارت (۳-۱۴) یا در یک بند دیگر تکمیل می کند [برای مثال: یکی بند موصولی (۳-۳) اسم اصلی (۳-۸) را در یک عبارت اسمی، پیرایش می کند].

یادآوری - یک بند پیرو معمولاً به خودی خود یک جمله نیست، ولیکن بخشی از یک جمله بزرگ تر است.

۱۸-۳

چارچوب طبقه بندی فرعی^۳

مجموعه ای از محدودیت هایی است، که نشان دهنده خواص موضوع های صرفی^۴ (۳-۱۹) است، که می توانند با فعل بیایند یا باید با فعل بیایند.

مثال: علی (/موضوع های صرفی /) کتاب می خواند (/موضوع صرفی /) امروز (/ادات /).

یادآوری - فاعل، مفعول غیر مستقیم و مفعول مستقیم زیرمجموعه های توابع دستوری (۳-۷) در یک جمله هستند و وابسته به فعل می باشند.

1-Span

2-Subordinate clause

3 - Subcategorization frame

4 - Syntactic argument

۱۹-۳

نهاد صرفی

عنصری است به لحاظ کاربردی ضروری، که مورد نیاز است و معنای خود را از هسته آن عبارت (۳-۱۴) یا از گره (۳-۱۲) که به آن وابسته است می‌گیرد (برای مثال: موضوع اسمی عبارت حرف اضافه یا فعل).

یادآوری - برای افعال و عبارات فعلی، نهادها شرکت کنندگان در روندی را شناسایی می‌کنند که با فعل به آن اشاره می‌شود. در برخی از چارچوب‌ها، نهادهای صرفی متمم نامیده می‌شوند.

۲۰-۳

نمودار صرفی^۱

نمودار

مقوله‌ای مرتبط از گره‌های صرفی (۳-۱۲) و لبه‌ها (۳-۶) است.

۲۱-۳

درخت صرفی^۲

نمودار صرفی (۳-۲۰) است، که در آن هر گره تنها دارای یک منشاء می‌باشد.

۲۲-۳

صرف^۳

شاخه‌ای از مطالعات زبانی که به مطالعه قواعد تشکیل سازه‌های بزرگتر از واژه در جمله سازی اختصاص دارد

۲۳-۳

گره پایانی

گره صرفی (۳-۱۲) است که یک صورت ساده و ازگانی (۳-۲۴) یا عنصر خالی درگیر در یک رابطه صرفی است.

1- Syntactic graph

2- Syntactic tree

3 -Syntax

شکل یک واژه خاص اعم از شکل نوشتاری یا آوایی آن است.

۴ فرآمدل SynAF

۱-۴ کلیات

حاشیه نویسی صرفی^۲ در پردازش زبان حداقل دو کاربرد به شرح زیر را دارد:
الف- برای نشان دادن سازواری زبان شناسی^۳، برای مثال: در عبارت اسمی (NP)، توالی ساختار بندی شده ای از مواردی را توصیف می کند که به صورت صرفی - نحوی حاشیه بندی شده اند (از جمله عناصر خالی یا آثار تولید شده به وسیله حرکات در سطح سازواری)، هم چنین ترکیبات ساخته شده از عناصر ناپیوسته و ب - برای نشان دادن روابط وابستگی، مانند روابط هسته - معرف و نیز از جمله روابط بین گروه های یکسان (مانند: روابط هسته - هسته بین اسامی در بدل ها، یا هماهنگی های اسمی در بعضی از صورتگرایی ها). اطلاعات وابستگی می تواند بین موارد حاشیه بندی شده به صورت صرفی - نحوی در یک عبارت وجود داشته باشند (صفت اصلاح کننده اسم اصلی در یک عبارت اسمی است) یا رابطه ای خاص بین ترکیبات صرفی در سطح بندها و جمله ها را توصیف کنند (یعنی «فاعل» فعل اصلی بند یا جمله یک عبارت اسمی باشد). ارتباط وابستگی هم چنین می تواند برای عناصر خالی نیز بیان شود (برای مثال: ضمیر مستور در زبان های عاشقانه، که نقشی دستوری می گیرد).

در نتیجه، همان گونه که در فرآمدل SynAF بیان شد، حاشیه نویسی صرفی باید منطبق با یک راهبرد چند لایه ای صرفی باشد که با حاشیه نویسی صرفی به هم پیوسته باشد، هم برای سازواری و هم برای وابستگی.

1-Word form

2- Syntactic Annotation

3- Linguistic constituency

۲-۴ فرآمدل SynAF

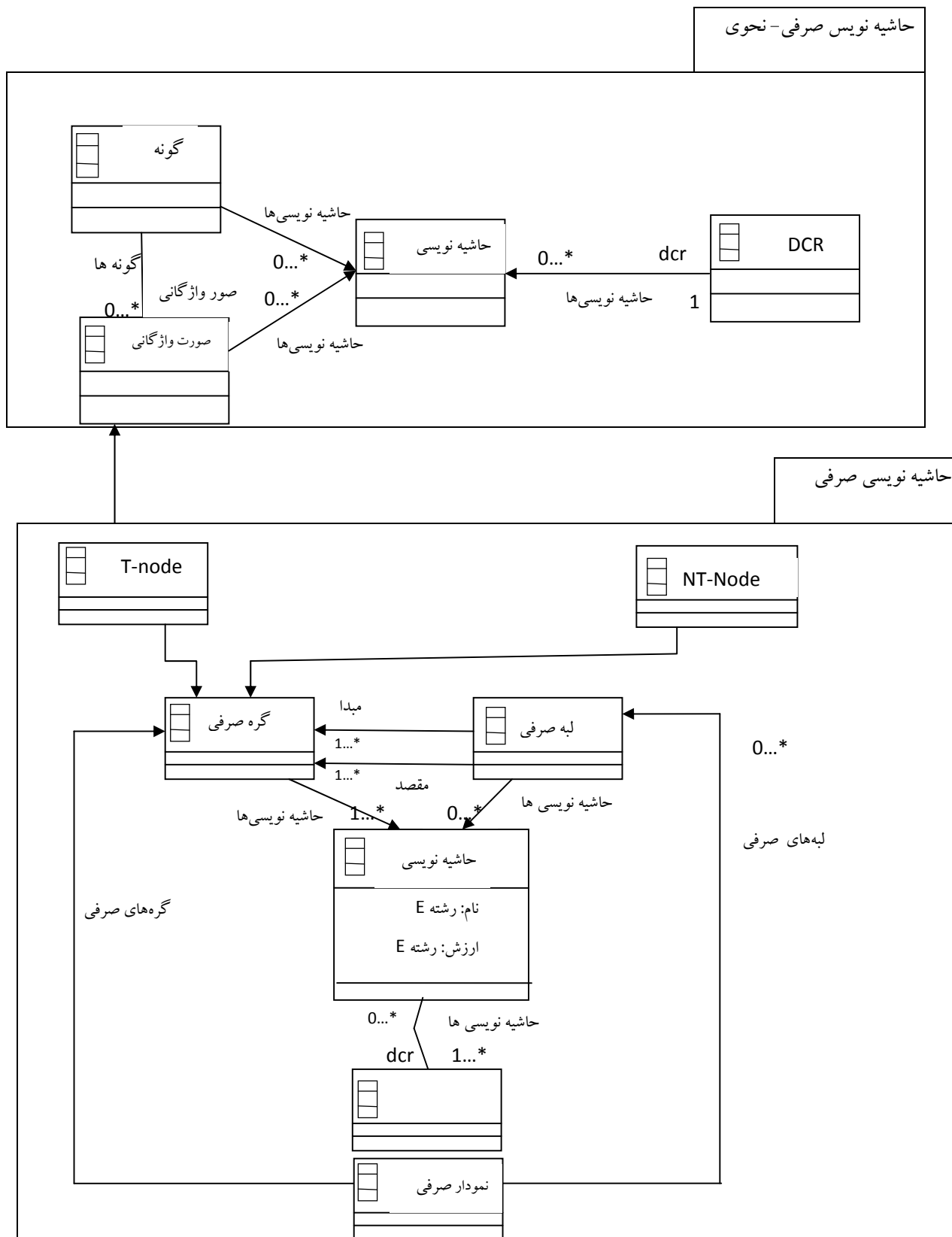
۱-۲-۴ رئوس کلی

فرآمدل SynAF به عنوان مجموعه‌ای از رده‌های زبان مدل سازی یکپارچه‌ای^۱ (UML) نشان داده می شود، که به وسیله جفت‌های صفت -ارزش UML تکمیل شده‌اند، که نمایانگر رده‌داده‌های صرفی وابسته می‌باشند. توصیفات متنی SynAF نسبت به آنچه که در نمودار UML گنجانده می شود، اطلاعات کامل تری در مورد رده‌ها، روابط و الحاقات SynAF را مشخص می‌کنند. طراحان باید انتخاب رده داده^۲ (DCS) را تعریف کنند، همان‌گونه که برای فرآیند انتخاب رده‌داده‌های SynAF مشخص شده (نگاه کنید به شکل ۱). دسته‌داده‌های داده شده در پیوست (الف) باید برای نشان دادن حاشیه نویسی صرفی استفاده شوند.

1- Unified Modeling Language

روشی برای تشخیص و شناسایی اشیا درون‌نظام و روابط بین آنها که به منظور طراحی و تحلیل نظام‌های شی‌گرا به کار می‌رود

2- Data category selection



شکل ۱- فرامدل SynAF (تقسیم بندی شده به وسیله MAF) چارچوب حاشیه نویسی فرامدل

۲-۲-۴

رده گره صرفی

رده گره صرفی، رده عمومی استقرایی برای هر دو رده گره‌های پایانی و رده از گره‌های غیر پایانی است. گره‌های صرفی می‌توانند در هر تعداد روابط صرفی مورد نیاز گنجانده شوند (نگاه کنید به ۳-۶، لبه‌های صرفی)

۳-۲-۴ رده $T\text{-Node}$ ^۱

رده $T\text{-Node}$ نشان دهنده گره‌های انتهایی درخت صرفی است، که صورت‌های واژگانی که به صورت صرفی - نحوی حاشیه گذاری شده اند و همین طور عناصر خالی، زمانی که مناسب هستند، را شامل می‌شود. گره‌های پایانی بر روی یک یا چند گستره تعریف می‌شوند. بیش از یک یا چند گستره (گستره‌های چندگانه می‌توانند مسئول سازه‌ها باشند). گره‌های پایانی با مقوله‌های صرفی حاشیه نویسی می‌شوند در سطح کلمه معتبر باشند.

۴-۲-۴ رده $NT\text{-Node}$ ^۲

رده $NT\text{-Node}$ نشان دهنده گره‌های غیر پایانی درخت نحوی می‌باشند. درختان صرفی به طور عمده عبارتند از گره‌های پایانی و گره‌های غیر پایانی شامل عناصر تهی وقتی مناسب هستند. گره‌های پایانی به یک گستره اشاره دارد. بنابراین، به موجب نمایش درخت صرفی، می‌توان گستره‌ها را نیز به جای گره‌های غیر پایانی به حساب آورد. این گره‌های غیر پایانی با مقوله‌های صرفی معتبر در سطح عبارت و بالاتر (بندی، جمله‌ای)، حاشیه نویسی می‌شوند.

۵-۲-۴

رده لبه صرفی

رده لبه صرفی نشان دهنده ارتباطی بین گره‌های صرفی (هم گره‌های پایانی وهم و غیر پایانی) است. برای مثال، رابطه وابستگی دوتایی است، متشکل از یک جفت از گره‌های منبع و مقصد، با یک یا چند نحو. خصوصاً لبه نحو می‌تواند به وسیله یک / نوع لبه صرفی^۳ / (به پیوست الف مراجعه کنید)، نحو شود، که مفهوم دامنه آن می‌تواند یکی از آنها باشد، ولیکن نه محدود به، / لبه صرفی اولیه^۴ /، / لبه صرفی ثانویه^۵ /.

1-Terminal Node

2- Non-terminal Node

3- Syntactic edge type

4 - Primary syntactic edge

5 - Secondary syntactic edge

۴-۲-۶ رده حاشیه نویسی

رده حاشیه نویسی نشان دهنده استفاده از اطلاعات صرفی برای داده‌های نحو شده SynAF، و هم چنین (نگاه کنید به شکل ۱) استفاده از اطلاعات صرفی- صرفی داده‌های صرف شده MAF است.

پیوست الف

(الزامی)

رده داده‌ها برای SynAF

الف-۱ کلیات

رده داده‌های زیر باید برای نشان دادن حاشیه نویسی صرفی در ترکیب با فرآمدل SynAF استفاده شوند. در صورت لزوم، برنامه‌های کاربردی خاص مجاز هستند رده داده‌های اضافی را تعریف کنند، که باید مطابق با ISO 12620 توصیف و در ثبت رده داده‌های ISO Cat آماده شوند.

الف-۲ رده داده‌های صرفی پایه

/ حاشیه نویسی /

تعریف - اطلاعاتی که به یک واژه، عبارت، بند، جمله، متن یا به رابطه میان آن‌ها اضافه شده است.

/ عمق حاشیه نویسی ^۱ /

دامنه مفهومی / تجزیه عمیق ^۲، / تجزیه سطحی ^۳، / پیوست افزایی ^۴ /

تعریف - سطح غنای اطلاعات که حاشیه نویسی توصیف می‌کند.

/ سبک حاشیه نویسی /

دامنه مفهومی / نشان‌گذاری درونه‌ای ^۵، / نشان‌گذاری مخلوط ^۶، / نشان‌گذاری توازن ^۷ /

تعریف - سبکی از حاشیه نویسی است.

1-Annotation depth

2- Deep parsing

3-Shallow parsing

4-Tagging

5-Embedded notation

6-Mixed notation

7-Stand-off notation

1 / نوع حاشیه نویسی

دامنه مفهومی / سازواری /، سازواری و وابستگی /، وابستگی /
تعریف - نوعی از حاشیه نویسی است

2 / واژه بست

تعریف - کلمه‌ای بدون تکیه است، که نمی تواند به تنهایی تلفظ شود و از لحاظ تلفظی وابسته به کلمات مجاور خود می باشد.

یادآوری - همه واژه‌بست‌ها در زبان فارسی از نوع پی‌بست هستند و به صورت پسوند ظاهر میشوند مانند «-ش» در «کتابش» که به معنای کتاب او است..

3 / سازواری

تعریف - ساز و کاری است، که اجازه تبدیل واژه‌ها به عبارات، و عبارات به عبارات یا بندهای بالاتر و بندها به جملات را می دهد.

یادآوری - تبدیل جملات به متن را معمولاً سازواری نمی نامند.

4 / سازواری و وابستگی

تعریف - اتحاد سازواری و وابستگی است.

5 / همجوار

تعریف - خاصیت یک واحد دستوری است، که مرزی را با دیگری به اشتراک می گذارد.

1 - Annotation type

2 - Clitics

3 - Constituency

4- Constituency and dependancy

5 - Contiguous

/ تجزیه عمیق^۱ /

تعریف - فرآیند رمزگشایی کامل بندها و روابط موجود در یک جمله می‌باشد.

/ وابستگی^۲ /

تعریف - ساز و کاری است، که اجازه اتصال واژه‌ها، یا در برخی صورتگرایی‌ها هم چنین عبارات و بندها، بر اساس تقسیم دوتایی هسته - وابسته و حاشیه نویسی محتمل نقش دستوری را می‌دهد.

/ نفی دوگانه^۳ /

تعریف - ساختاری است متشکل از دو شکل منفی در بندی یکسان.

/ نشان گذاری درونه ای^۴ /

تعریف - حاشیه نویسی است، که به متن اضافه می‌شود.

یادآوری - سازمان اصلی متن اصلاح می‌شود.

/ پی بست^۵ - رجوع شود به : / واژه بست /

تعریف - واژه بستنی است که به واژه پیش از خود وابسته است.

/ اول /

تعریف - پیش از هر چیز با توجه به نظمی خاص می‌آید.

1- Deep parsing

2 - Dependency

3 - Double negative

4 - Embedded notation

5 - Enclitic

/ نشانه گذاری مخلوط^۱ /

تعریف- حاشیه نویسی به سبک پیوندی که در آن درونه ای و توازن مخلوط هستند.

/ حاشیه نویسی صرفی - نحوی^۲ / - رجوع شود به : / حاشیه نویسی /

تعریف- حاشیه نویسی به نحو واژه‌ها و نوع آنها مرتبط می‌باشد.

/ نفی /

تعریف- ساختاری است، که رد یا انکار بخش یا تمام معنای یک جمله، واژه یا عبارت را بیان می‌کند.

یادآوری- برای افاده مفهوم منفی در فارسی می‌توان از نشانه‌های نفی هر دو حوزه صرف و نحو استفاده کرد:

مثال: تصمیم او درست نیست (حوزه نحو). تصمیم او نادرست است (حوزه صرف).

/ بعدی^۳ /

تعریف- بلافاصله بعد است.

/ لبه صرفی اولیه^۴ /

تعریف- لبه پیش فرضی است که رابطه سازواری، به وجود آمدن از یک سازه و خاتمه یافتن در یک جزء از آن سازه را بیان می‌دارد.

/ محمول^۵ /

تعریف- عبارت یا واژه‌ای در یک بند است که عبارتی را درباره موضوع آن بند بیان می‌کند. بنابراین، اکثر بندها می‌توانند به دوبخش موضوع و محمول تقسیم شوند، که در آن محمول نقشی است که در باره موضوع توضیح می‌دهد.

1 - Mixed notation

2- Morpho syntactic annotation

3- Next

4- Primary syntactic edge

5- Predicate

یادآوری - برای مثال : "کیان توپ را زد" به این صورت دیده می شود، که ("کیان" موضوع است و به همراه عبارت محمول ("توپ را زد") آمده است.

/ پیشین^۱ /

تعریف - بلافاصله پیش از

/ پیش بست^۲ - رجوع شود به : واژه بست

تعریف - واژه بستنی است که بستگی به کلمه بعدی دارد.

یادآوری - برای مثال : حرف تعریف "the" در "the boy"

/ انتشار^۳ /

تعریف - عمل گسترش یک خاصیت زبانی از یک واحد دستوری به واحدی دیگر است.

/ لبه صرفی ثانویه^۴ /

تعریف - یک لبه غیر مستقیم برای بیان سازواری صرفی است . این لبه ها می توانند برای بیان رابطه بین هسته ای و هم خانواده ای از وابسته محذوف آن مورد استفاده قرار گیرند.

یادآوری - مثال : در جمله «آرش را دیدم، اما پس از آن مستقیم به خانه برگشتم»، «من» ممکن است به عنوان فاعل صریح و روشن بند اول به کار رود، تحت تسلط یک لبه صرفی اولیه باشد، ولیکن در بند دوم، یک لبه صرفی ثانویه اضافی که منجر به «من» می شود، می تواند این موضوع را روشن کند، که در بند دوم نیز «من» فاعل است، بدون این که یکی از بخش های آشکار آن بند باشد، که به وسیله لبه های اولیه فراگرفته شده اند. این وسیله در برخی از صورتگرایی ها برای جلوگیری از معرفی عناصر خالی که جایگزین حاملین پنهان نقش دستوری هستند، استفاده می شود.

1- Previous

2-Proclitic

3-Propagation

4-Secondary syntactic edge

/ تجزیه سطحی^۱ /

تعریف- روند شناسایی قطعه‌ها در یک جمله است.

/ حاشیه نویسی توازن^۲ /

تعریف- حاشیه نویسی توضیحی است در خارج از واحدهای دستوری ثبت شده و آنچه به این واحدها اشاره می‌کند.

یادآوری - سازمان اصلی متن دست نخورده نگه داشته می‌شود.

/ حاشیه نویسی صرفی / - رجوع شود به : / حاشیه نویسی /

تعریف- حاشیه نویسی‌ای که سازواری و / یا وابستگی را توصیف می‌کند.

یادآوری - حاشیه نویسی صرفی به صورت مستقیم با معنای کلام سر و کار ندارد.

/ مشخصه صرفی^۳ /

تعریف- ویژگی‌ای که در شرح نحو یک زبان استفاده می‌شود.

/ نوع لبه صرفی^۴ /

دامنه مفهومی / لبه صرفی اولیه /، / لبه صرفی ثانویه /

تعریف- لبه صرفی را با توجه به نقش آن در نمایش صرفی مشخص می‌کند.

/ محدودیت صرفی^۵ /

-
- 1-Shallow parsing
 - 2-Standoff notation
 - 3-Syntactic feature
 - 4-Syntactic edge type
 - 5-Syntactic restriction

تعریف- قاعده‌ای است که آنچه را نحو در زبان خاصی مجاز دانسته، محدود می‌کند.

/ برچسب زنی^۱ /

تعریف- روند شرح نوع کلمه برای هر واژه است.

/ نوع استفهامی /

تعریف- خاصیت یک بند که با یک واژه سئوالی شروع می‌شود.

یادآوری - در انگلیسی جمله او کیست؟، یک سؤال نوع استفهامی است .

/ نوع بله- نه /

تعریف- خاصیت یک بند که در آن تنها یک پاسخ یا موقعیت مثبت یا منفی امکان پذیر است.

یادآوری - در فارسی، «آیا شما می‌آیید؟» یک سوال نوع بله- نه ای است.

الف-۳ رده‌داده‌های مربوط به ساختار سازواری

/ قطعه صفت^۲ / - رجوع شود به : / قطعه /

تعریف - قطعه‌ای که یک صفت واژه اصلی آن است.

/ گروه صفتی^۳ / - رجوع شود به : / عبارت /

تعریف- عبارتی است، که یک صفت واژه اصلی آن است .

/ قطعه افزودنی^۴ / - رجوع شود به : / قطعه /

1- Tagging

2- Adjective chunk

3- Adjective phrase

4- Adposition chunk

تعریف - قطعه‌ای است که به وسیله یک یا چند افزودنی آغاز می‌شود، که لزوماً به هم پیوسته و در یک انتهای قطعه نیستند.

/ عبارت افزودنی^۱ /

تعریف - عبارتی است، که به وسیله یک یا چند افزودنی آغاز می‌شود و حاوی متممی هم چون گروه اسمی است .
یادآوری - افزودنی‌ها لزوماً به هم پیوسته و در یک انتهای قطعه نیستند.

/ قطعه قید^۲ / - رجوع شود به : / قطعه /

تعریف - قطعه‌ای است که هسته آن یک قید می‌باشد.

/ عبارت قیدی^۳ / - رجوع شود به : / عبارت /

تعریف - عبارتی است که هسته آن یک قید است .

/ قطعه^۴ / - رجوع شود به : / واحد دستور زبان /

تعریف - یک سری واژه‌های غیر تصریفی، که عموماً حاوی بیش از یک واژه می‌باشند.

یادآوری - یک قطعه نمی‌تواند حاوی یک زیر ساخت باشد. قطعه اغلب شبیه به یک عبارت و بیشتر اوقات استمراری است.

/ بند^۵ / - رجوع شود به : / واحد دستور زبان /

تعریف - واحدی از سازمان دستوری، کوچک تر یا برابر با جمله ولیکن بزرگ تر از عبارت و واژه ، و به طور کلی حاوی گزاره خود است .

یادآوری - طبقه بندی سنتی بندها شامل بندهای اصلی (مستقل یا مافوق) و تابع (یا وابسته) است ، برای مثال : پسر رسید (بند اصلی) پس از آن که باران آغاز شده (بند فرعی). بند مجاز است یک جمله کامل باشد، برای مثال : "آن‌ها آمدند". بند مجاز است در بردارندهٔ فرو بند^۶ باشد.

1- Adposition phrase

2-Adverb chunk

3-Adverb phrase

4- Chunk

5- Clause

6--Sub- clause

1 / عبارت مقایسه‌ای / - رجوع شود به : / عبارت /

تعریف- عبارتی است که معنای تطبیقی را بیان می‌کند.

یادآوری- در زبان فارسی، به صورت صرفی (برای مثال: بزرگ تر) برای بیان مقایسه وجود دارد.

2 / عبارت هم‌پایه / - رجوع شود به : / عبارت /

تعریف - عبارتی است، که هم‌پایگی را بیان می‌کند.

3 / بند اخباری / - رجوع شود به : / بند /

تعریف- بندی است که به بیان یک گفته اشاره و دارای یک ارزش حقیقی مشروط می‌باشد.

یادآوری- معمولاً این عبارت در تقابل با بندهای پرسشی و امری محسوب می‌شود.

4 / واحد دستوری /

تعریف- اصطلاحی است که به یک واژه، عبارت، بند یا جمله اشاره می‌کند.

5 / بند امری / - رجوع شود به : / بند /

تعریف- بندی است که به بیان فرمان اشاره دارد.

یادآوری- معمولاً این عبارت در تقابل با بندهای پرسشی و اخباری محسوب می‌شود.

6 / بند پرسشی / - رجوع شود به : / بند /

تعریف- بندی است که به بیان یک سوال اشاره می‌کند.

1- Comparative phrase

2-Coordinate phrase

3- Declarative clause

4- Grammatical unit

5- Imperative clause

6- Interrogative clause

یادآوری - معمولاً، این عبارت در تقابل با بندهای اخباری و امری محسوب می شود. برای مثال: در زبان فارسی، "تو کی هستی؟".

/ تکه اسمی^۱ / - رجوع شود به : / قطعه /
تعریف - قطعه‌ای است که هسته آن یک اسم است.

/ گروه اسمی^۲ / - رجوع شود به : / عبارت /
تعریف - عبارتی است که هسته آن یک اسم است.

/ عبارت^۳ / - رجوع شود به : / واحد دستوری /
تعریف - عنصر ساختاری است که در اطراف هسته تشکیل شده، تعیین کننده خواص دستوری آن می باشد و از صفر، یک یا چند واژه و/ یا عبارات دیگر که بدون ساختار فاعل - گزاره مرسوم در بندها، ساخته شده است.

یادآوری - عبارت ممکن است زیر سازه‌هایی را احاطه کند. به طور سنتی عبارت به عنوان بخشی از یک سلسله مراتب ساختاری محسوب که بین بند و کلمه افتاده است. معمولاً با توجه به هسته، چندین نوع را می توان از هم تمییز داد، برای مثال : عبارت قیدی، عبارت صفتی، و غیره.

/ قطعه پسایند^۴ / - رجوع شود به : / قطعه /
تعریف - قطعه ای است که کلمه اصلی آن پسایند است.

/ عبارت پسایندی^۵ / - رجوع شود به : / عبارت /
تعریف - عبارتی است که کلمه اصلی آن پسایند است.

1-Noun chunk

2-Noun phrase

3-Phrase

4-Postposition chunk

5- Postposition phrase

/ قطعه حرف اضافه^۱ / - رجوع شود به : / قطعه /
تعریف - قطعه‌ای است که کلمه اصلی آن حرف اضافه است .

/ گروه حرف اضافه^۲ / - رجوع شود به : / عبارت /
تعریف - عبارتی است که کلمه اصلی آن یک حرف اضافه می‌باشد.

/ حرف اضافه گروه فعلی^۳ / - رجوع شود به : / عبارت /
تعریف - عبارت فعلی است که با یک حرف اضافه شروع می‌شود.

/ بند موصولی^۴ / - رجوع شود به : / بند /
تعریف - بندی است که به عنوان صفت برای گروه اسمی به کار می‌رود و با ضمیر موصولی شروع می‌شود که ممکن است محذوف باشد.

یادآوری - در زبان انگلیسی، بند موصولی با ضمیر موصولی شروع می‌شود، مانند "که". جملات نسبی ممکن است تحدیدی باشند، زیرمجموعه ای از اسمی که آن‌ها بسط می‌دهند را مشخص کنند یا غیرتحدیدی باشند و صرفاً یک ویژگی را اضافه کنند، برای مثال: "مردانی که در حال مبارزه بودند شجاع بودند" (و آن‌هایی که مبارزه نمی‌کردند چنین نبودند) در مقابل "مردان، که در حال مبارزه بودند، شجاع بودند" (همه مردان شجاع بودند، و اتفاقاً، آن‌ها در حال مبارزه بودند).

/ جمله / - رجوع شود به : / واحد دستور زبان /
تعریف - سازمان دستوری است متشکل از یک بند اصلی و تمام بندهای تابع و بندهای تابع آن بندها به همین ترتیب تا الاخر.

یادآوری - جمله ممکن است به انواع ساده و مرکب رده‌بندی شود، یعنی این که شامل یک واحد فاعل-گزاره یا بیش از یکی باشد.

/ عبارت عالی^۵ / - رجوع شود به : / عبارت /

-
- 1- Preposition chunk
 - 2- Preposition phrase
 - 3- Preposition verb phrase
 - 4- Relative clause
 - 5- Superlative clause

تعریف- عبارتی است برای بیان معنی عالی.

یادآوری - در زبان فارسی، شکل صرفی (برای مثال : بزرگترین) برای بیان عالی وجود دارد.

/ سازه صرفی^۱ /

تعریف- یک واحدی دستوری است که بخشی از یک واحد بزرگتر دستوری را تشکیل می‌دهد، یک سازه از آن واحد بعدی است. اگر دو واحد به طور مستقیم به وسیله یک لبه سازواری صرفی به هم وصل شوند، یک سازه مستقیم و در غیر این صورت یک سازه غیر مستقیم است.

/ هسته فعل^۲ - رجوع شود به : / قطعه /

تعریف - قطعه ای است متشکل از فعلی که گسسته شده و احتمالاً به همراه واژه بستوار خود می‌باشد.

/ عبارت فعلی / - رجوع شود به : / عبارت /

تعریف- گروهی است که هسته آن یک فعل است.

الف-۴ رده‌داده‌های مربوط به وابستگی

/ معرف صفت^۳ / - رجوع شود به : / معرف صفت /

تعریف- رابطه ای است که در آن تعدیل از طریق یک صفت انجام می‌گیرد.

/ ادات /

تعریف- واحد دستوری انتخابی یا ثانویه است، که ممکن است بدون این که باقیمانده ساختار تحت تاثیر قرار گیرد، حذف شود.

یادآوری - قیدها معمولاً ادات هستند، همان طور که در "پرهام توپ را زد" به جای "دیروز پرهام توپ را زد"، که هر دو دستوری می‌باشند.

1- Syntactic constituent

2- Verb nucleus

3- Adjective modifier

۱ / معرف قید / - رجوع شود به : / معرف /
تعریف- رابطه‌ای است که در آن تعدیل شده یک قید است.

۲ / بدل شده /
تعریف- خاصیت در بدل بودن است.

۳ / بدل شدن /
تعریف- رابطه بین واحدهای زبانی است که مرجع یکسان (یا مشابه) و کاربرد دستوری یکسان در بند یکسان را با یک واحد که واحدهای دیگر را بسط می‌دهد، به شراکت می‌گذارد.
یادآوری- برای مثال : "بهبود، آرایشگر، وارد شد".

۴ / صفت /
تعریف- ارتباطی است که به صفت یا اسم می‌پردازد زمانی که به عنوان تعدیل کننده کلمه اصلی عبارت اسمی به کار می‌روند.

۵ / فعل کمکی /
تعریف- رابطه بین فعل تابع و فعل اصلی است.

۶ / روابط مقایسه‌ای / - رجوع شود به : / ارتباط /
تعریف- رابطه‌ای است که معنای تطبیقی را بیان می‌کند.

۱ / متمم ساز /

-
- 1 - Adverb modifier
 - 2- Apposed
 - 3 - Apposition
 - 4 - Attribute
 - 5- Auxilliary
 - 6-Comparative relation

تعریف- رابطه بین حرف ربط وابسته و فعل است که یک جمله درونه‌ای از نوع متمم را نشان می‌دهد.

همپایگی /^۲

تعریف - فرایند یا نتیجه ابعاد رابطه برابر بین دو یا چند سازه که به لحاظ نحوی یکسان هستند

هم پایه‌ساز /^۳

تعریف - حروف و نشانه‌های همپایگی را تعدیل می‌کنند

یادآوری- عموماً هم‌پایه‌ساز یک حرف ربط هم پایه است.

مفعول مستقیم /^۴

تعریف- رابطه بین فعل و نهاد آن است که در آن، نهاد به وسیله یک بدل وساطت نمی‌شود یا با حالتی غیر فاعلی نشان داده نمی‌شود.

یادآوری- در جمله‌ی "مرد به پسر کتابی داد"، عبارت "کتاب" مفعول مستقیم است، در حالی که بسیاری از صورت‌گرایها "پسر" را به عنوان مفعول غیرمستقیم به حساب می‌آورند که مسئول حالت فرضی (و به صورت تاریخی مشاهده شده) غیر فاعلی (حالت مفعولی) می‌باشد.

حالت اضافه /^۵ - رجوع شود به : / رابطه /

تعریف- رابطه‌ای است که معمولاً ملکیت را بیان می‌کند یا در غیر اینصورت رابطه وابسته اسمی یک اسم معرف را بیان می‌کند.

هسته /^۶

1 - Complementizer

2 - Coordination

3- Coordinator

4 - Direct object

5 - Genitive

6 - Head

تعریف- کلمه مرکزی یک عبارت است که خواص دستوری و توزیعی آن را تعیین می‌کند.

/ معرف^۱ /

تعریف- عنصر زبانی (برای مثال: صورت واژگانی) که آغاز یک گروه صرفی را نشان می‌دهد.

یادآوری- برای مثال: یک حرف اضافه معرف یک عبارت اسمی است.

/ هم‌نهی^۲ /

تعریف- رابطه‌ای است که در آن دو واحد زبانی در کنار هم قرار می‌گیرند.

/ همپایه شده از چپ^۳ /

تعریف- عضوی از بند است که در سمت چپ هم پایه ساز قرار می‌گیرد.

/ پیراینده^۴ /

تعریف- رابطه‌ای یک واحد دستوری با واحدی دیگر است که آن واحد اولی به آن بستگی دارد و دیگری را گسترش می‌دهد.

یادآوری- در "درخت بزرگ در باغ"، هم "بزرگ" و هم "در باغ"، "درخت" را مشخص می‌کنند.

/ معرف اسم^۵ / - رجوع شود به: / معرف /

تعریف- رابطه‌ای است که در آن تعدیل شده یک اسم است.

/ ساختار توصیفی پس اسمی^۶ / - رجوع شود به: / پیراینده /

1- Introducer

2- Juxtaposition

3- Left coordinated

4- Modifier

5- Noun modifier

6- Post nominal modifier

تعریف- تعدیل کننده اسمی است، که در آن معرف پس از اسم می آید.

/ ساختار توصیفی پیش اسمی^۱ / - رجوع شود به : / پیراینده /

تعریف- تعدیل کننده اسمی است، که در آن معرف پیش از اسم می آید.

/ معرف حرف اضافه^۲ / - رجوع شود به : / پیراینده /

تعریف- رابطه‌ای است که در آن مشخص شده حرف اضافه است.

/ ارتباط /

تعریف- پیوند تشریح شده بین دو یا چند واحد دستوری است

/ رابطه موصولی^۳ / - رجوع شود به : / ارتباط /، / معرف اسمی /

تعریف- رابطه‌ای است که بند موصولی را به گروه اسمی پیوند می دهد که به آن اشاره دارد.

/ هم پایه شده از راست^۴ /

تعریف- عضوی از بند است که در سمت راست هم پایه ساز قرار می گیرد.

/ هسته ساختار^۵ /

تعریف- بولی^۶ است که نشان می دهد جزء مورد نظر، هسته صرفی است یا نه.

/ فاعل /

تعریف- ارتباط بین عبارت و فعل است که شخص یا چیزی را نشان می دهد که چیزی برای آن بیان می شود.

/ رابطه عالی^۱ / - رجوع شود به : / رابطه /

1 - Prenominal modifier

2 - Preposition modifier

3- Relative relation

4- Right coordinated

5- Structure head

6- Boolean

سیستم بولی بر این اساس است که چیزی یا می تواند درست باشد یا اشتباه و نمی تواند هر دوی آنها باشد

تعریف- رابطه‌ای است که معنایی عالی را بیان می‌کند.

/ موضوع نحوی ۲ /

تعریف- موضوعی از ساختاری صرفی است.

/ عملکرد نحوی ۳ /

تعریف- رابطه‌ای است بین شکل زبانی و دیگر بخش‌های سیستم زبانی که در آن، این شکل زبانی استفاده شده است.

یادآوری- برای مثال: فاعل.

/ هسته صرفی ۱ /

تعریف- عنصری است که توزیع و خواص دستوری واحدی را تعیین می‌کند که در بردارنده خود واحد و وابسته‌های آن است.

/ متمم فعل ۴ /

تعریف- ارتباط بین عبارت و فعل، که در آن عبارت داده شده برای فعل مرکزی نیست.
یادآوری- بر عکس آنچه در مورد مفعول مستقیم گفته شد.

/ معرف فعل ۵ / - رجوع شود به : / معرف /

تعریف- رابطه‌ای است که نشان‌دهنده تعدیل فعل است.

1-Superlative relation

2-Syntactic argument

3- Syntactic function

4- Syntactic head

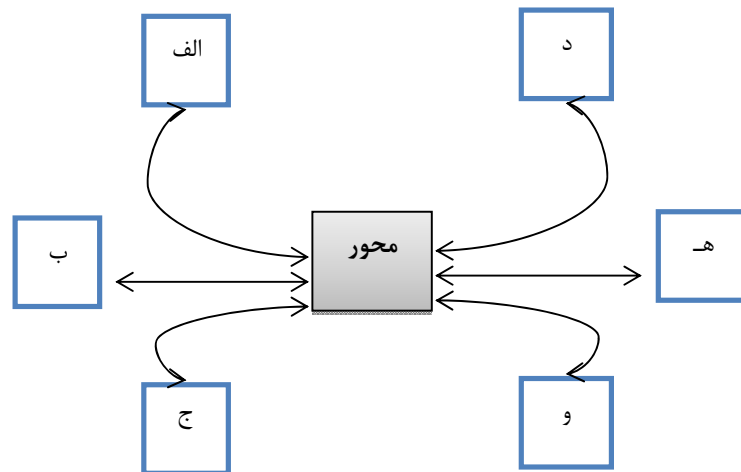
5 -Verb modifier

پیوست ب

(اطلاعاتی)

ارتباط به چارچوب حاشیه نویسی زبانی

برای اجرای SynAF (LAF؛ چهارچوب زبانی حاشیه نویسی) استاندارد ISO 24612 باید مورد توجه قرار گیرد. LAF چارچوب کلی برای ارائه حاشیه نویسی را فراهم می کند، همان گونه که در پیوست کتابشناسی بندهای [14]-[15]-[16] و Rommary 2004, 2006 [14]-[15]-[16] شرح داده شده است. توسعه آن بر پایه عمل مشترک و هم‌گرایی روش‌های حاشیه نویسی زبانی در طول ۲۰-۱۵ سال گذشته، ایجاد شده است. هسته چارچوب، مشخصات مدل انتزاعی است که برای حاشیه نویسی به وسیله قالب محوری نمونه سازی شده است، که حاشیه نویسی به منظور تبادل به داخل و خارج از آن ترسیم شده است.



شکل ب-۱- LAF به صورت قالب محوری

شکل ب- ۱ اصول LAF برای شش کاربر متفاوت و قالب حاشیه نویسی‌هایی را نشان می‌دهد (از «الف» تا «و») نام گذاری شده اند، که به دو تنظیم کننده برای هر طرحواره نیاز دارد- یکی به داخل و یکی به خارج از قالب محوری، و توسط طراح طرح ارائه می‌شود. بنابراین، حداکثر تعداد تنظیم کننده‌ها در طرح‌ها $2n$ ، در حالی که $n^2 - n$ تنظیم کننده‌های دوطرفه بدون محور است.

- طرح حاشیه نویسی برای تنظیم کننده‌ها در سطح محوری باید مطابق با مدل انتزاعی باشد (یا از طریق نقشه برداری تبدیل شود) که متشکل است از (۱) یک ساختار ارجاعی برای تداعی حاشیه نویسی‌های توازن با داده‌های اولیه، که به عنوان یک نمودار جهت دار نمونه سازی شده است و (۲) یک ارائه ساختار مشخصه برای محتوای حاشیه نویسی. بدین ترتیب حاشیه نویسی یک نمودار جهت داری را تشکیل می‌دهد که به n بعدی داده‌های اولیه اشاره دارد و همین طور دیگر حاشیه نویسی‌ها که در آن‌ها گره‌ها با دیگر ساختارهای مشخصه‌ای حاشیه نویسی می‌شوند که محتوای حاشیه نویسی را فراهم می‌کنند. به صورت رسمی، LAF عبارتند از:

- یک مدل داده‌ای برای حاشیه نویسی که بر اساس نمودار جهت دار به شکل زیر تعریف شده است: یک نمودار از حاشیه نویسی‌های G یک رده از رئوس $V(G)$ (راس مترادف گره است) و یک رده از لبه‌های $E(G)$ است. رئوس و لبه‌ها ممکن است با یک یا چند ویژگی حاشیه نویسی شوند. یکی ویژگی متشکل از یک چهار بخشی (G', VE, K, V) است که در آن G' نمودار، VE راس یا لبه در G' ، K اسم ویژگی و V ارزش مشخصه است. - تقسیم بندی اساسی داده‌های اولیه که لبه‌های بین گره‌های مجازی را تعریف می‌کند که بین هر "نویسه" در داده‌های اولیه قرار گرفته است، و در آن هر کارکتر چنین تعریف شده که دنباله بایتی به هم پیوسته به طولی مشخص (برای متن، پیش فرض 16 - UTF است) باشد. نمودار حاصله G به عنوان یک نمودار لبه G' شناخته می‌شود که گره‌های آن، لبه‌های G می‌باشند و به عنوان گره‌های برگ ("sink") عمل می‌کنند. این گره‌ها پایه‌ای را برای یک حاشیه نویسی یا چند لایه حاشیه نویسی فراهم می‌آورند. تقسیم بندی چندگانه می‌تواند بر روی داده‌های اولیه تعریف شده باشد و حاشیه نویسی‌های متعدد، ممکن است به همان تقسیم بندی اشاره کنند.

- ترتیب بندی مدل داده‌ها که یکی از آنها به عنوان محور تعیین شده ؛

- روش‌هایی برای بهره‌برداری کردن از مدل داده‌ها.

توجه داشته باشید که LAF مختصه‌هایی برای رده‌بندی محتوای حاشیه نویسی فراهم نمی‌آورد (یعنی حاشیه نویسی پدیده‌های زبانی مرتبط را توصیف می‌کند)، که برای آن استاندارد سازی امری زیرکانه‌تر است. بنابراین، معماری LAF به یک ثبت دسته د/دها (DCR) اشاره دارد که شامل عناصر داده‌های از پیش تعریف شده و

طرح‌هایی هستند که ممکن است به همراه ابزاری برای مشخص نمودن مقوله‌های جدید و تعدیل مقوله‌های موجود در حاشیه نویسی‌ها به صورت مستقیم استفاده شود. (به [15][14] Ide and Romary, 2004 مراجعه کنید).

پیوست پ

(اطلاعاتی)

کتابنامہ

1. ISO 639-1:2002, *Codes for the representation of names of languages — Part 1: Alpha-2 code*
2. ISO 639-2:1998, *Code for the representation of names of languages — Part 2: Alpha-3 code*
3. ISO 639-3:2007, *Codes for the representation of names of languages — Part 3: Alpha-3 code for comprehensive coverage of languages*
4. ISO/IEC 10646-1:2000, *Information technology — Universal Multiple-Octet Coded Character Set (UCS) — Part 1: Architecture and Basic Multilingual Plane*
5. ISO/IEC 11179-3:2003, *Information technology — Metadata registries (MDR) — Part 3: Registry metamodel and basic attributes (MDR3)*
6. ISO 24610-1:2006, *Language resource management — Feature structures — Part 1: Feature structure representation*
7. ISO 24612, *Language resource management — Linguistic annotation framework (LAF)*
8. ISO 24613, *Language resource management — Lexical markup framework (LMF)*
9. ABEILLÉ, A. (ed.) *Building and Using Syntactically Annotated Corpora*. Kluwer, Dordrecht, 2001
10. [10] ABEILLÉ, A., HANSEN-SCHIRRA, S. and USZKOREIT, H. (eds.), *Proceedings of the 4th International Workshop on Linguistically Interpreted Corpora (LINC-03)*, 2003
11. CALZOLARI, N., MCNAUGHT, J. and ZAMPOLLI, A. (eds). *EAGLES: Introduction, 1996*.
12. [12] FRANCOPOULO, G., DECLERCK, T., SORNLERLAMVANICH, V., de la CLERGERIE, E. and MONACHINI, M. 2008. Data Category Registry: Morpho-syntactic and Syntactic profiles, *LREC Workshop on use and usage of language resource-related standards*
13. IDE, N. and ROMARY, L. A common framework for syntactic annotation. In *Proceedings of the 39th Annual Meeting on Association for Computational Linguistics*, Toulouse, France, July 6-11, 2001.
14. IDE, N. and ROMARY, L. A Registry of Standard Data Categories for Linguistic Annotation. *Proceedings of the Fourth Language Resources and Evaluation Conference (LREC)*, Lisbon, 2004, pp. 135-139
15. IDE, N. and ROMARY, L. International Standard for a Linguistic Annotation Framework. *Journal of Natural Language Engineering*, **10**:3-4, 2004, pp. 211-225
16. IDE, N. and ROMARY, L. Representing Linguistic Corpora and Their Annotations. *Proceedings of the Fifth Language Resources and Evaluation Conference (LREC)*, Genoa, Italy, 2006
17. IDE, N. *GrAF: A Graph-based Format for Linguistic Annotations*. Proceedings of the LAW Workshop at ACL 2007, Prague, 2007
18. MONTEMAGNI, F. *et al.* Building the Italian Syntactic-Semantic Treebank. In: *Building and using Parsed Corpora* (ed. Abeillé, A.), Language and Speech series, Kluwer, Dordrecht, 2003
19. RUMBAUGH, J., JACOBSON, I. and BOOCH, G. *The Unified Modeling Language Reference Manual*, 2nd edition. Addison Wesley, 2004