

عددشکن فارسی*

سید مهدی سمائی (استادیار پژوهشگاه علوم و فناوری اطلاعات ایران)

مقوله‌بندی پدیده‌ها بر اساس کمیّت یکی از ابزارهای شناخت است. اعداد واحدهای شناختی محاسبه کمیّت‌اند. انگشت‌های دست و پا ابتدایی‌ترین وسیله محاسبه بوده‌اند. از این رو، پایه واژگانی شمارش در بیشتر زبان‌ها اعداد ۵ و ۱۰ و ۲۰ است (Gvozdanovic 2006, p. 736). به جز بعضی زبان‌های بومیان استرالیا، که در آنها نظام اعداد وجود ندارد و تمایزهای کمیّ با صور دستوری بیان می‌شود، اغلب زبان‌ها نظام عددی دارند. این نظام‌ها، بر اثر پیشرفت‌های علم ستاره‌شناسی و هندسه و به اقتضای نیاز در محاسبات تقویم‌نگاری و مبادلات تجاری، به تدریج پدید آمده‌اند (Ibid, p. 737). اعداد ترکیبی و پیچیده را معمولاً با عمل جمع و گاهی ضرب می‌سازند. نشان دادن اعداد مرکب با تفریق هم داشته‌ایم اما نادر است^۱. زبان‌هایی نظیر زبان یوروبا^۲، که در آنها، هم عمل

* عددشکن فارسی طرحی بوده در تقطیع اعداد که در پژوهشگاه علوم و فناوری اطلاعات ایران اجرا شده است. مهندس مهرداد نوروزی برنامه‌نویس طرح بوده و نسترن یزدانی قواعد صوری نظام اعداد فارسی را برای برنامه‌نویسی آماده کرده است. این مقاله خلاصه‌ای از طرح مذکور است که، در آن، قواعد ترکیب نام‌های اعداد اصلی آمده است.

(۱) مثلاً، در خط لاتینی، عدد ۹ را به صورت IX نمایش می‌دهند که در آن X (مفروق‌منه) = ۱۰ و I (مفروق) = ۱ است و IX نمایش ۹-۱=۱۰ است. مفروق در سمت چپ مفروق‌منه قرار می‌گیرد. در حالی که XI (I در سمت راست X) نمایش ۱۱=۱۰+۱ است.

(۲) Yoruba زبانی بومی که هفده میلیون نفر در نیجریه و عده‌ای هم در بنین و توگو به آن تکلم می‌کنند.

جمع و هم تفریق برای ساختن اعداد ترکیبی به کار رود، نیز نادرند. (Hurford 1992, p. 132)
در زبان فارسی نیز، همانند اغلب زبان‌ها، نظام عددگویی وجود دارد. همچنین نظامی
برای عرضه اعداد به خط فارسی هست که به آن نظام عددنویسی می‌گویند. در نظام
عددنویسی، رقم بسیط با رابطه‌های جانشینی^۳ و همنشینی^۴ ارزش مشخصی در درون
رقم مرکب پیدا می‌کند (نظیر ارزش هر رقم بسیط در عدد ۲۳۱۰۵۹۶۸). در نظام عددگویی
زبان فارسی چند گروه وجود دارد که اعضای آنها با قواعد معینی با هم ترکیب می‌شوند. نام
اعضای این گروه‌ها را به خط فارسی می‌نویسند (نظیر بیست و سه میلیون، صد و پنج هزار،
نهصد و شصت و هشت). بنابراین، در نظام اعداد زبان فارسی نیز، مانند دیگر زبان‌ها،
از قواعد صرفی و صرفی و نحوی برای بیان اعداد پیچیده استفاده می‌شود.^۵

سابقه بررسی نظام اعداد فارسی

در باره قواعد ترکیب اعداد در زبان فارسی، هرمز میلانیان (۱۳۵۱) و نگارنده این مقاله
(۱۳۷۷ و ۱۳۸۴) تحقیق کرده‌ایم. میلانیان به رقم‌نویسی (گونه نوشتاری بیان عدد) و
رقم‌گویی (گونه گفتاری بیان عدد) اشاره می‌کند و نظام‌های رقم‌نویسی را محدود و جهانی
می‌داند. او تفاوت نظام عددی در زبان‌های گوناگون را در شیوه بیان رقم‌گویی می‌داند.
میلانیان نظام رقم‌نویسی فارسی را که اصلیتی هندی دارد به چهار گروه و چند عضو
منفرد تقسیم می‌کند. (جدول ۱)

ترکیب اعضای این گروه‌ها با یکدیگر به دو روش انجام می‌شود:

الف- روش جمع

در این روش، عدد کوچک‌تر پس از عدد بزرگ‌تر می‌آید و حرف عطف /و/ رابط آنها
می‌شود: بیست و چهار.

ب- روش ضرب

در این روش، عدد کوچک‌تر پیش از عدد بزرگ‌تر می‌آید و در آن ضرب می‌شود: سه‌هزار.

3) paradigmatic relations 4) syntagmatic relations

۵) با عددشکن فارسی می‌توان ترکیب‌های دستوری نظام عددگویی فارسی را تشخیص داد و معادل آن را
در نظام عددنویسی فارسی عرضه کرد.

جدول ۱. گروه‌های چندعضوی دستگاه عددگوئی فارسی

الف	ب	ج	د	ه	و
یک	ده	بیست	صد	هزار	میلیون
دو	یازده	سی	دویست		
سه	دوازده	چهل	سیصد		
چهار	سیزده	پنجاه	چهارصد		
پنج	چهارده	شصت	پانصد		
شش	پانزده	هفتاد	ششصد		
هفت	شانزده	هشتاد	هفتصد		
هشت	هفده	نود	هشتصد		
نه	هجده		نهصد		
	نوزده				

امکانات اعداد ترکیبی را می‌توان با ادغام روش‌های الف و ب افزایش داد. مثلاً با چهل و دو و سه‌هزار می‌توان عدد بزرگ‌تر سه‌هزار و چهل و دو را ساخت. نگارنده، در رسالهٔ دکتری خود (۱۳۷۷) و در مقالهٔ «گروه اعداد در ترجمهٔ ماشینی» (۱۳۸۴)، قواعد ترکیب اعداد را از دیدگاه کاربرد در رایانه بررسی کرده است. از آنجا که مباحث نظری آن دو منبع در این مقاله به صورتی عملی و در چارچوب نرم‌افزار عددشکن عرضه شده است، نیازی به عرضهٔ جزئیات آنها در این قسمت نیست.

قواعد اعداد ترکیبی در عددشکن

قواعدی که در طرح عددشکن برای امکانات ترکیب اعداد عرضه شده است دو تفاوت عمده با کار میلانین دارد: اولاً این قواعد نتیجهٔ مطالعه‌ای کاربردی است که با هدف پردازش زبان^۶ و از این دیدگاه انجام شده که به کار ساختن ماشینی هوشمند بیاید؛ ثانیاً برای هر یک از امکانات ترکیب اعداد قاعده‌ای مشخص پیشنهاد شده و به بیان کلیات بسنده نشده است. قواعد مطرّد (استثنانپذیر) و صریح است و درستی و دقت آن در رایانه آزمایش شده است.

6) language processing

اعداد اصلی شامل چهار طبقه اصلی و چند طبقه منفردند^۷

جدول ۲. قواعد گروه اعداد اصلی

مثال	قاعده ترکیب طبقات*	
بیست و یک	ج + حرف عطف + الف	۱
سیصد و دو	د + حرف عطف + الف	۲
چهارصد و دوازده	د + حرف عطف + ب	۳
نهد و سی	د + حرف عطف + ج	۴
هزار و پنج	ه + حرف عطف + الف	۵
هزار و سیزده	ه + حرف عطف + ب	۶
هزار و چهل	ه + حرف عطف + ج	۷
هزار و سیصد	ه + حرف عطف + د	۸
میلیون و سه	و + حرف عطف + الف	۹
میلیون و سیزده	و + حرف عطف + ب	۱۰
میلیون و سی	و + حرف عطف + ج	۱۱
میلیون و سیصد	و + حرف عطف + د	۱۲
میلیون و هزار	و + حرف عطف + ه	۱۳
سیصد و بیست و یک	د + حرف عطف + ج + حرف عطف + الف	۱۴
هزار و شصت و دو	ه + حرف عطف + ج + حرف عطف + الف	۱۵
هزار و صد و یک	ه + حرف عطف + د + حرف عطف + الف	۱۶
هزار و پانصد و شانزده	ه + حرف عطف + د + حرف عطف + ب	۱۷
هزار و هشتصد و هفتاد	ه + حرف عطف + د + حرف عطف + ج	۱۸
هزار و نهصد و هشتاد و یک	ه + حرف عطف + د + حرف عطف + ج + حرف عطف + الف	۱۹
میلیون و پنجاه و شش	و + حرف عطف + ج + حرف عطف + الف	۲۰
میلیون و سیصد و نه	و + حرف عطف + د + حرف عطف + الف	۲۱
میلیون و سیصد و سی	و + حرف عطف + د + حرف عطف + ب	۲۲
میلیون و هزار و دو	و + حرف عطف + ه + حرف عطف + الف	۲۳

* مرجع حروف ابجد در این ستون جدول ۱ است.

(۷) چهار طبقه اصلی همان چهار گروهی هستند که در جدول ۱ آمده است. طبقات منفرد شامل هزار و میلیون هستند که میلیارد/بیلیون (۱۰^۹) و تریلیون (۱۰^{۱۲}) را هم می‌توان به آن افزود.

مثال	قاعده ترکیب طبقات	
میلیون و هزار و دوازده	و + حرف عطف + ه + حرف عطف + ب	۲۴
میلیون و هزار و سیصد	و + حرف عطف + ه + حرف عطف + د	۲۵
میلیون و هزار و چهل و دو	و + حرف عطف + ه + حرف عطف + ج + حرف عطف + الف	۲۶
میلیون و هزار و صد و دو	و + حرف عطف + ه + حرف عطف + د + حرف عطف + الف	۲۷
میلیون هزار و نهصد و سی ^۸	و + حرف عطف + ه + حرف عطف + د + حرف عطف + ج	۲۸
دو هزار	الف + ه	۲۹
دوازده هزار	ب + ه	۳۰
چهل هزار	ج + ه	۳۱
چهارصد هزار	د + ه	۳۲
دو میلیون	الف + ه	۳۳
سیزده میلیون	ب + و	۳۴
نود میلیون	ج + و	۳۵
هشتصد میلیون	د + و	۳۶
پنجاه و دو هزار	ج + حرف عطف + الف + ه	۳۷
سیصد و یک هزار	د + حرف عطف + الف + ه	۳۸
هزار و دوازده هزار	د + حرف عطف + ب + ه	۳۹
ششصد و پنجاه هزار	د + حرف عطف + ج + ه	۴۰
سیصد و سی و دو هزار	د + حرف عطف + ج + حرف عطف + الف + ه	۴۱
سی و دو هزار	ج + حرف عطف + الف + ه	۴۲
سیصد و سه میلیون	د + حرف عطف + الف + و	۴۳
دویست و دوازده میلیون	د + حرف عطف + ب + و	۴۴
ششصد و سی میلیون	د + حرف عطف + ج + و	۴۵
پانصد و سی و دو میلیون	د + حرف عطف + ج + حرف عطف + الف + و	۴۶

قاعده‌های شماره ۱ تا ۴۶ در مرحله آخر برای برنامه‌نویسی و کاربرد در رایانه با حروف قراردادی لاتینی رمزگذاری شده‌اند.

(۸) پیش از قاعده‌های شماره ۹ تا ۱۳ و ۲۰ تا ۲۸ یکی از اعضای گروه الف قرار می‌گیرد.

منابع

سمائی، سید مهدی (۱۳۷۷)، واژگان در دستورسنج (رساله دکتری زبان‌شناسی)، دانشگاه تهران، تهران.
_____ (۱۳۸۴)، «گروه اعداد در ترجمه ماشینی»، فصلنامه علوم و فناوری اطلاعات، سال ۲۱، ش ۱،
ص ۲۱-۳۳.

میلانیان، هرمز (۱۳۵۱)، «دستگاه عدد در زبان فارسی»، مجموعه خطابه‌های سومین کنگره تحقیقات ایرانی، ج ۱،
به کوشش محمد روشن، بنیاد فرهنگ ایران، تهران، ص ۶۳۱-۶۴۵.

Gvozdanovic, J. (2006), "Numeral Systems: Their Structures and Development", *Encyclopedia of Language and Linguistics*, Keith Brown (ed.), Elsevier, pp. 736-739.

Hurford, James (1992), "Numerals", *International Encyclopedia of Linguistics*, William Bright (ed.),
Oxford University Press, pp. 131-132.

