**Stroke-level Modeling for Chinese Characters (or 字 zì)**

There are 32 (or 33) types of strokes. 8 or so types are essential, which are used in other composite types as parts. 永 (yŏng) 字八法(zìbāfǎ) is a well-known phrase indicating all the eight (八) essential types are represented in the 永character (字).

[32 types are for teaching purposes. With variants such as brought in with bend and slant, more are listed with <https://commons.wikimedia.org/wiki/Category:CJK_strokes_(SVG)> for printing purposes.]

* Six out of the eight types are standalone (e.g., 撇 and 捺 used in 八, and 点 the dot used 永, 字, and 点 itself).
* 折 (indicating a 90 degree turn) is used as a joint in a composite stroke to connect 横and a subsequent 竖segment, or vice versa. As shown in the chart, 折is highlighted at the turn in the central stroke in 永, connecting a short 横 segment with a long 竖segment to form a composite stroke 横折钩![A black background with a black square

  Description automatically generated with medium confidence](data:image/png;base64,iVBORw0KGgoAAAANSUhEUgAAACEAAAAhBAMAAAClyt9cAAAAAXNSR0IArs4c6QAAAARnQU1BAACxjwv8YQUAAAAPUExURQAAAAAAAAAAAAAAAAAAAE8O540AAAAFdFJOUwD/OJjRflpWOAAAAAlwSFlzAAAh1QAAIdUBBJy0nQAAAEJJREFUKM9joBQICqLzBcFCjBAuI5CnACRhIkCuAEQc3SBMEWZBTNsGUoSFCDcPsIgAugpBNBEGTBFGLCIM5AIGBgAWMwMhw3mYuAAAAABJRU5ErkJggg==).
* 钩 can only be appended to other strokes, going down or going left only; except going up when attached with bend or slant strokes.

|  |  |
| --- | --- |
| A blue and black symbol with white text  Description automatically generated | Essential types – standalone  点——dot， 即tiny dash。  横——horizontal，即rightward stroke。  竖——vertical，即downward stroke。  提——rise或rising，即flick up and rightwards。尾部变尖锐（taper）。  捺——press down或right-falling，即falling rightwards，flattening at the bottom。  撇——throw away或left-falling，即falling leftwards，with slight curve，not very curved。 |
| Essential types – Joint/attached  折——break或turning，即90 degree turn，going down or going right only。  钩——hook，即appended to other strokes，going down or going left only。 | Variant styles  弯——bend，即usually concave on the left。  斜——slanty，即usually concave on the right。 |

A StrokeEnum type is created to define strokes, using names (in pinyin) as listed in the following table. An abbreviation is coined up with initial letters of each pinyin term, such as HZG for 横折钩![A black background with a black square

Description automatically generated with medium confidence](data:image/png;base64,iVBORw0KGgoAAAANSUhEUgAAACEAAAAhBAMAAAClyt9cAAAAAXNSR0IArs4c6QAAAARnQU1BAACxjwv8YQUAAAAPUExURQAAAAAAAAAAAAAAAAAAAE8O540AAAAFdFJOUwD/OJjRflpWOAAAAAlwSFlzAAAh1QAAIdUBBJy0nQAAAEJJREFUKM9joBQICqLzBcFCjBAuI5CnACRhIkCuAEQc3SBMEWZBTNsGUoSFCDcPsIgAugpBNBEGTBFGLCIM5AIGBgAWMwMhw3mYuAAAAABJRU5ErkJggg==) (heng zhe gou).

Only 11 types are downloaded (from CJK commons site?) for the Chn123 GUI prototype, with a different coding scheme, {"d", "h", "v", "t", "p", "u", "hv", "vj", "vaj"} for {D, H, S, …, SWG}, etc.



[Gou (钩 or hook) is always an “end effect”; while Zhe (折) may be a Heng or Shu segment.]

<https://www.facebook.com/xueyejiaoyu/posts/%E6%B1%89%E5%AD%97%E7%AC%94%E7%94%BB%E5%90%8D%E7%A7%B0%E8%A1%A8%E6%94%B6%E8%97%8F%E7%89%88%E5%B0%8F%E5%AD%A6%E7%94%9F%E5%BF%85%E5%AD%A6%E9%80%82%E5%90%88%E4%B8%80%E4%BA%8C%E4%B8%89%E5%B9%B4%E7%BA%A7%E5%AD%A6%E7%94%9F%E5%AD%A6%E5%8F%B6%E6%95%99%E8%82%B2%E7%BD%91/113116250615869/>

**public** **enum** StrokeEnum {

***HENG***, ***SHU***, ***PIE***, ***NA***, ***DIAN***, ***TI***, // essential

***SHU\_GOU***, ***HENG\_ZHE***, ***HENG\_ZHE\_GOU***, ***HENG\_GOU***,

***HENG\_PIE***, ***SHU\_ZHE***, ***SHU\_TI***, // intermediate

***PIE\_TI***, ***PIE\_DIAN***, ***XIE\_GOU***, ***SHU\_WAN\_GOU***,

***SHU\_ZHE\_ZHE\_GOU***, ***HENG\_WAN\_GOU***, ***HENG\_PIE\_WANG\_GOU***, // advanced

***HENG\_ZHE\_TI***, ***WAN\_GOU***, ***WO\_GOU***, ***HENG\_ZHE\_WAN\_GOU***,

***HENG\_ZHE\_ZHE\_ZHE\_GOU***, ***SHU\_WAAN***, ***HENG\_ZHE\_WAN***,

***HENG\_ZHE\_ZHE\_PIE***, ***SHU\_ZHE\_PIE***, ***HENG\_XIE\_GOU***,

***SHU\_ZHE\_ZHE***, ***HENG\_ZHE\_ZHE***,***HENG\_ZHE\_ZHE\_ZHE***;

// 32 in all ==> 33 w/ HWG

33

[DIAN, HENG, HENG\_GOU, HENG\_PIE, HENG\_PIE\_WANG\_GOU, HENG\_WAN\_GOU, HENG\_XIE\_GOU, HENG\_ZHE, HENG\_ZHE\_GOU, HENG\_ZHE\_TI, HENG\_ZHE\_WAN, HENG\_ZHE\_WAN\_GOU, HENG\_ZHE\_ZHE, HENG\_ZHE\_ZHE\_PIE, HENG\_ZHE\_ZHE\_ZHE, HENG\_ZHE\_ZHE\_ZHE\_GOU, NA, PIE, PIE\_DIAN, PIE\_TI, SHU, SHU\_GOU, SHU\_TI, SHU\_WAAN, SHU\_WAN\_GOU, SHU\_ZHE, SHU\_ZHE\_PIE, SHU\_ZHE\_ZHE, SHU\_ZHE\_ZHE\_GOU, TI, WAN\_GOU, WO\_GOU, XIE\_GOU]

H S P N D T

SG HZ HZG HG HP SZ ST PT PD XG SWG SZZG HWG HPWG HZT WG WG HZWG HZZZG SW HZW HZZP SZP HXG SZZ HZZ HZZZ

A stroke knowledge model is defined in the stroke package, with

* A stroke.model package includes an interface type IStroke and a few abstract types (EssentialStroke and CompositeStroke in model package).
* The StrokeEnum codifies the 33 stroke types…
* An EndPoint is used to define an EssentialStroke, each of which has exactly two end points.
* In the stroke package, a concrete HengStroke class is defined, which extends the EssentialStroke class, with more details to be added.
* A StrokeSequence class is defined to hold strokes in the right order for a character (or Zi).
* Classes for thinning and testing are included in a stroke.thinning package.
* The stroke.image package has a couple of utility classes (and a DAO class for now).

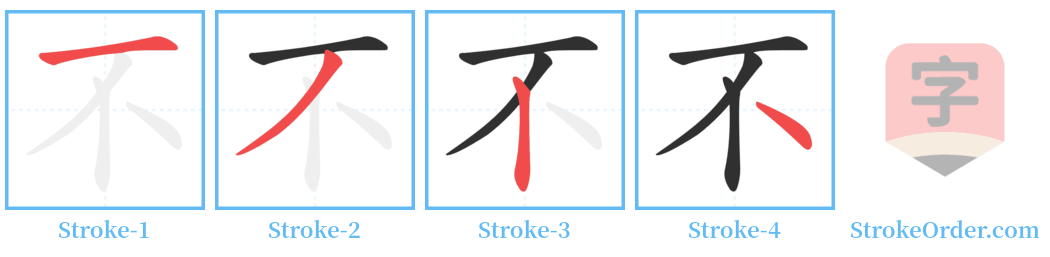
Several test drivers are included in the stroke.test package. Among them,

* The MiZiGeLabel class can generate a panel with a grid box with a 米 (mi) shape to help place strokes in a square, as automated using the StrokeInPlace class. The stroke.image.StrokeInPlace class provides utility functions to match the stroke svg images (stored in the images/strokes/kaiti folder) into the grid box.
* The LearnFromStrokeSeqDemo class demonstrates using pixel scanning of a stroke sequencing image accessible on Hanziwu.com. Three heng-strokes are determined by the leftmost and rightmost points of red-colored segments from a sequence for 三. And the MiZiGeLabel class shows the “learning” results.

To dos:

* Implement classes similar to HengStroke (formally) for other stroke types used in characters for the numbers: H, S, P, N, D, HZ, SW, SWG, HZWG, as shown below.
* Extends capability of the LearnFromStrokeSeqDemo to learn stroke sequence knowledge for 一 through 十, as well as a few more characters不, 人, 大, 天.
* Store the sequence (in a text file or DB) and test with visualizing them in the grid box created using the MiZiGeLabel class, and rendering to the anatomy lab.

|  |  |  |
| --- | --- | --- |
| For characters 一 二三四五六七八九 十, and 不, 人, 大, 天 | | |
| H 一二三 十 五 六 四 七 不 大 天  S 十 五 四 不  P 八 六 九 不 人大 | N 八 六 四 不 人大  D 六  HZ 五 四 | SW 四  SWG 七  HZWG 九 |



[On <https://www.hanziwu.com/>, copy the character such as 不 from the list to the search box, which will present the sequence.]

[Nov 17]

Created classes for HSPN stroke types, and a test bank with 一三 (learned from hanziwu.com) and 人, 大, 天, 王 (with handset end points).

A screenshot of a computer screen

Description automatically generatedA screen shot of a computer screen

Description automatically generated

Pie and na strokes need to have the knowledge of curved contour to trace the stroke better.

[Java QuadCurve2D or CubicCurve2D can be used to render the curved strokes.]

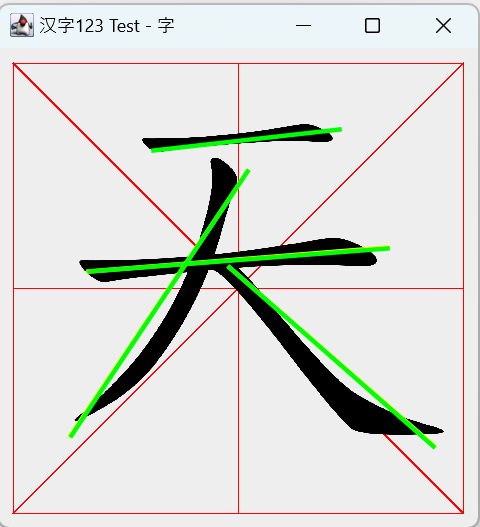
The StrokeSeq class or the ChnChar class that holding a stroke seq should be able to

* Know the total stroke count.
* Types of strokes used.
* Pick the stroke by index or relative position.

Individual strokes don’t typically carry specific meaning.

Meaning carrying modules and radicals are where “story”-based knowledge will be embedded.

[Nov 20]

A screenshot of a computer screen

Description automatically generatedA screenshot of a computer

Description automatically generated

Learned 天 with all strokes taken as HENG. End points for PIE need to be swapped.

A screenshot of a computer screen

Description automatically generatedA screenshot of a computer screen

Description automatically generated

changed the last stroke to a dian…

SHU strokes are hard-coded and adjusted based on the look in the box.

[Dec 6] Coming back after Thanksgiving break and wrapping up with classes

A screenshot of a computer screen

Description automatically generated 🡺 A screenshot of a computer screen

Description automatically generated

A screenshot of a computer screen

Description automatically generatedAdded curved stroke (using QuadCurve2D.Double)

Trying to generate JSON for ChnChars:

HENG(0.301,0.235)-->(0.719,0.199)

HENG(0.316,0.48)-->(0.694,0.454)

HENG(0.112,0.735)-->(0.949,0.709)

{"chnChar":"三", "unicode":"\u4e09","strokes":

["strokeType":"HENG","points":[{"x":0.301,"y":0.235},{"x":0.719,"y":0.199}],

"strokeType":"HENG","points":[{"x":0.316,"y":0.48},{"x":0.694,"y":0.454}],

"strokeType":"HENG","points":[{"x":0.112,"y":0.735},{"x":0.949,"y":0.709}]

]}

{"chnChar":"一", "unicode":"\u4e00","strokes":

["strokeType":"HENG","points":[{"x":0.102,"y":0.495},{"x":0.923,"y":0.49}]

]}

{"chnChar":"大", "unicode":"\u5927","strokes":

["strokeType":"HENG","points":[{"x":0.199,"y":0.434},{"x":0.806,"y":0.378}],

"strokeType":"PIE","points":[{"x":0.51,"y":0.112},{"x":0.133,"y":0.847}],

"strokeType":"NA","points":[{"x":0.469,"y":0.459},{"x":0.939,"y":0.847}]

]}

{"chnChar":"三", "unicode":"\u4e09","strokes":

["strokeType":"HENG","points":[{"x":0.301,"y":0.235},{"x":0.719,"y":0.199}],

"strokeType":"HENG","points":[{"x":0.316,"y":0.48},{"x":0.694,"y":0.454}],

"strokeType":"HENG","points":[{"x":0.112,"y":0.735},{"x":0.949,"y":0.709}]

]}

{"chnChar":"天", "unicode":"\u5929","strokes":

["strokeType":"HENG","points":[{"x":0.311,"y":0.194},{"x":0.724,"y":0.148}],

"strokeType":"HENG","points":[{"x":0.168,"y":0.464},{"x":0.832,"y":0.413}],

"strokeType":"PIE","points":[{"x":0.52,"y":0.24},{"x":0.128,"y":0.827}],

"strokeType":"NA","points":[{"x":0.48,"y":0.454},{"x":0.934,"y":0.852}]

]}

{"chnChar":"人", "unicode":"\u4eba","strokes":

["strokeType":"PIE","points":[{"x":0.551,"y":0.189},{"x":0.066,"y":0.791}],

"strokeType":"NA","points":[{"x":0.454,"y":0.403},{"x":0.964,"y":0.796}]

]}

{"chnChar":"王", "unicode":"\u738b","strokes":

["strokeType":"HENG","points":[{"x":0.281,"y":0.224},{"x":0.74,"y":0.184}],

"strokeType":"HENG","points":[{"x":0.316,"y":0.495},{"x":0.704,"y":0.459}],

"strokeType":"SHU","points":[{"x":0.5,"y":0.224},{"x":0.5,"y":0.75}],

"strokeType":"HENG","points":[{"x":0.107,"y":0.77},{"x":0.929,"y":0.765}]

]}

{"chnChar":"二", "unicode":"\u4e8c","strokes":

["strokeType":"HENG","points":[{"x":0.281,"y":0.301},{"x":0.73,"y":0.255}],

"strokeType":"HENG","points":[{"x":0.112,"y":0.668},{"x":0.913,"y":0.653}]

]}

{"chnChar":"不", "unicode":"\u4e0d","strokes":

["strokeType":"HENG","points":[{"x":0.173,"y":0.224},{"x":0.867,"y":0.184}],

"strokeType":"PIE","points":[{"x":0.633,"y":0.25},{"x":0.117,"y":0.724}],

"strokeType":"SHU","points":[{"x":0.55,"y":0.374},{"x":0.55,"y":0.91}],

"strokeType":"DIAN","points":[{"x":0.617,"y":0.469},{"x":0.888,"y":0.663}]

]}

Story annotations:

王, meaning king.

The top HENG indicates heaven, and the bottom HENG for land. The middle HENG indicates human or people. The SHU stroke indicates that “the king” is the only one that can connect the three: with the mandate from Heaven to rule all People on the Land.

八, meaning eight, which is most divisible in all 10 numbers: 8=2x2x2

The PIE and NA strokes are separated and ready to be divided.

六, meaning six, which reuses the character 八 as its lower part. Indicating it’s an even number.

人, meaning human or person

The PIE and NA strokes are joined near top, making a human figure standing straight up.

SAI/Computing Conference 2025 - references

Influence of stroke-order learning on Chinese character recognition among Chinese as a second language learners

Author links open overlay panelHui Sun a, Tianlin Wang b, Miao Yu a

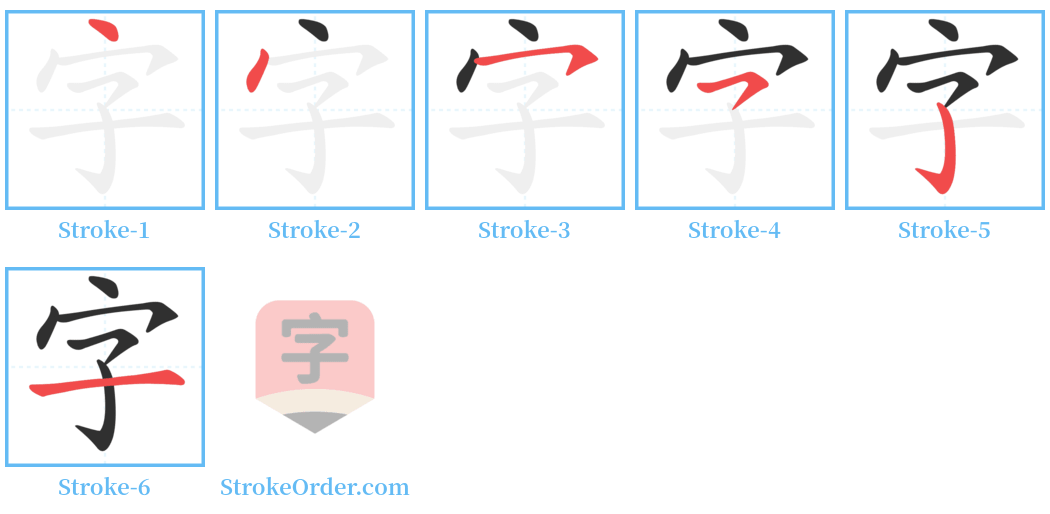
<https://www.sciencedirect.com/science/article/pii/S0346251X24000770?via%3Dihub>

Interference effects of radical markings and stroke order animations on Chinese character learning among L2 learners

[[](https://loop.frontiersin.org/people/1493306)Fengyun Hou](https://loop.frontiersin.org/people/1493306)[[Xin Jiang
](https://loop.frontiersin.org/people/1355855)Xin Jiang](https://loop.frontiersin.org/people/1355855)

<https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.783613/full>

To-dos

* Replace placeholder code (lines 50-100) in LearnFromStrokeSeqDemo that deal with all strokes as HENG to specific types (such asDIAN, PIE, NA).
* Retrieve stroke-wise illustration images for CHN123 characters (next page) from Hanziwu.com and save them in the lmzis folder.
* Modify code to deal with characters with more than 5 strokes (with more than one line of grids) so that all stroke boxes can be handled.
* Include the new characters learned in the StrokeTestSeqBank class for testing.
  + Optionally, save character stroke info as a JSON file using the JsonGenerator class.

CHN123 Characters

String lessonFile = "ChnChar LessonNo SeqNoInLesson\r\n" +

"一 1 1\r\n" +

"二 1 2\r\n" +

"三 1 3\r\n" +

"一 2 -1\r\n" +

"二 2 -2\r\n" +

"三 2 -3\r\n" +

"四 2 4\r\n" +

"五 2 5\r\n" +

"六 2 6\r\n" +

"七 2 7\r\n" +

"八 2 8\r\n" +

"九 2 9\r\n" +

"十 2 10\r\n" +

"日 3 1\r\n" +

"月 3 2\r\n" +

"星 3 3\r\n" +

"期 3 4\r\n" +

"生 3 5\r\n" +

"其 3 6\r\n" +

"金 4 1\r\n" +

"木 4 2\r\n" +

"水 4 3\r\n" +

"火 4 4\r\n" +

"土 4 5\r\n" +

"人 5 1\r\n" +

"大 5 2\r\n" +

"天 5 3\r\n" +

"地 5 4\r\n" +

"也 5 5\r\n" +

"辶 5 99\r\n" +

"道 5 99\r\n" +

"今 6 1\r\n" +

"昨 6 2\r\n" +

"是 6 4\r\n" +

"口 7 1\r\n" +

"舌 7 2\r\n" +

"目 7 3\r\n" +

"耳 7 4\r\n" +

"鼻 7 5\r\n" +

"自 7 6\r\n" +

"左 8 1\r\n" +

"右 8 2\r\n" +

"手 8 3\r\n" +

"足 8 4\r\n" +

"又 8 5\r\n" +

"东 9 1\r\n" +

"西 9 2\r\n" +

"南 9 3\r\n" +

"北 9 4\r\n" +

"中 9 5\r\n" +

"父 10 1\r\n" +

"母 10 2\r\n" +

"子 10 3\r\n" +

"女 10 4\r\n" +

"儿 10 5\r\n" +

"肉 11 1\r\n" +

"背 11 2\r\n" +

"腿 11 3\r\n" +

"脚 11 4\r\n" +

"臂 11 5\r\n" +

"腰 11 6\r\n" +

"要 11 99\r\n" +

"首 12 1\r\n" +

"面 12 2\r\n" +

"头 12 3\r\n" +

"兄 13 1\r\n" +

"弟 13 2\r\n" +

"姐 13 3\r\n" +

"妹 13 4\r\n" +

"祖 13 5\r\n" +

"礻 13 99\r\n" +

"示 13 99\r\n" +

"且 13 99\r\n" +

"未 13 99\r\n" +

"有 13 99\r\n" +

"上 14 1\r\n" +

"下 14 2\r\n" +

"午 14 3\r\n" +

"早 14 4\r\n" +

"晚 14 5\r\n" +

"晨 14 6\r\n" +

"小 15 1\r\n" +

"时 15 2\r\n" +

"分 15 3\r\n" +

"钟 15 4\r\n" +

"秒 15 5\r\n" +

"寸 15 99\r\n" +

"刀 15 99\r\n" +

"江 16 1\r\n" +

"河 16 2\r\n" +

"湖 16 3\r\n" +

"海 16 4\r\n" +

"山 16 5\r\n" +

"川 16 6\r\n" +

"氵 16 99\r\n" +

"胡 16 99\r\n" +

"每 16 99\r\n" +

"古 16 99\r\n" +

"红 17 1\r\n" +

"黄 17 2\r\n" +

"黑 17 3\r\n" +

"白 17 4\r\n" +

"蓝 17 5\r\n" +

"绿 17 6\r\n" +

"田 17 99\r\n" +

"灬 17 99\r\n" +

"艹 17 99\r\n" +

"百 18 1\r\n" +

"千 18 2\r\n" +

"万 18 3\r\n" +

"年 18 4\r\n" +

"半 18 5\r\n" +

"好 20 1\r\n" +

"坏 20 2\r\n" +

"多 20 3\r\n" +

"少 20 4\r\n" +

"不 20 5\r\n" +

"学 21 1\r\n" +

"习 21 2\r\n" +

"汉 21 3\r\n" +

"字 21 4\r\n" +

"之 21 99\r\n" +

"而 21 99\r\n" +

"说 22 1\r\n" +

"普 22 2\r\n" +

"通 22 3\r\n" +

"话 22 4\r\n" +

"言 22 99\r\n" +

"了 23 1\r\n" +

"解 23 2\r\n" +

"中 23 -3\r\n" +

"国 23 4\r\n" +

"文 23 5\r\n" +

"化 23 6\r\n" +

"不 24 -1\r\n" +

"用 24 2\r\n" +

"死 24 3\r\n" +

"记 24 4\r\n" +

"己 24 99\r\n" +

"我 25 1\r\n" +

"你 25 2\r\n" +

"他 25 3\r\n" +

"她 25 4\r\n" +

"它 25 5\r\n" +

"们 25 6\r\n" +

"";