## 4. 100 surnames [25 marks]

When the Mongol Emperor Kublai Khan initiated the Yuan dynasty (1271–1368) in China, he commissioned Lama ´Gro-mgon Chos-rgyal ´Phags-paa to create a unified script to write all the major languages under his rule. Although the resulting system (now called ´Phags-pa) never caught on beyond official use, some classic Chinese texts survive in a ´Phags-pa version.

The Băijiāxìng (Hundred Surnames) is a Song Dynasty (960–1279) poem listing over 400 classical Chinese family names. Although originally written in Chinese characters, during the Yuan dynasty this poem was written in ´Phags-pa characters as well, as shown below.

The illustration shows two consecutive pages of the Băijiāxìng Měnggǔwén ("The Hundred Surnames in Mongol Script"), from a 1340 manuscript.

Below are twenty lines (9-28) from the Yuan-era Băijiāxìng, with some names missing (identified by row number and column letter). The two pages given above correspond to a portion of the poem below. Your task is to figure out which portion of this poem the pages represent, and use this to figure out what the missing names must be.
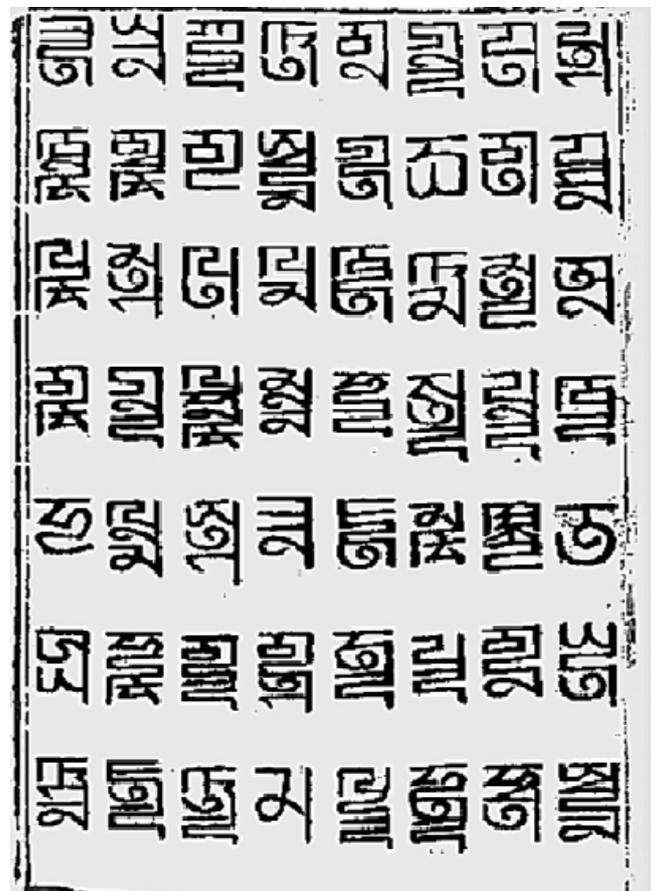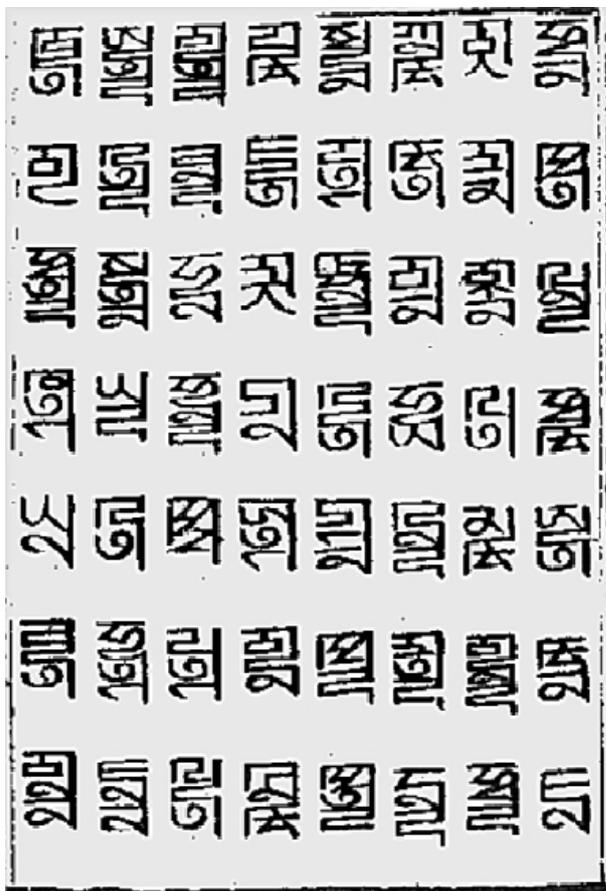
|	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
9	Fi	Lem	Drxim	Sĭa,	Lue	Ho	Yi	Thang
10	Dxing	'In	Lo	Pi,	Haŭ	'U	'An	Srang
11	Yaŭ	Yĭu	Sri	Fu,	Bue	Pen	Dzi	Khang
12	U	Yĭu	Ngĭŭan	Pu,	Ku	Mung	Bing	Hŏang
13	[13a]	Fu	Sring	Taĭ,	Dam	Sung	[13g]	Bang
14	Xĭung	Ki	[14c]	Khĭu,	[14e]	Trĭu	Tung	Lĭang
15	Du	[15b]	Lam	[15d],	Zi	[15f]	[15g]	Gĭang
16	Kĭa	Lu	Lxiŭ	Ngue,	Kĭang	Dung	[16g]	Kŭaŭ
17	[17a]	Sring	Lim	Xĭaŭ,	Trung	Zĭu	Khiŭ	Laŭ
18	Kaŭ	[18b]	Tshaĭ	Den,	Fan	Hu	[18g]	Faŭ
19	Ngĭu	Wan	Tri	Ko,	[19e]	Kŏan	Lu	Maŭ
20	Kĭing	[20b]	[20c]	Wu,	Kan	Xĭaĭ	'Ing	Tsung
21	Ting	Sĭŭan	Pue	Dxing,	'Ĭu	Sren	Hang	Hung
22	Paŭ	Trĭu	[22c]	Sri,	Tshue	Kĭi	Nriŭ	Kĭung
23	Dring	Xĭi	Xĭing	[23d],	Bue	Lĭu	Ngĭung	'Ung
24	Sĭun	Yang	[24c]	Xĭue,	Trin	Khĭu	Kĭa	Fung
25	Nyue	Yi	Drĭu	Kin,	Ki	Ping	Mue	Zĭung
26	Tsing	Dŏan	Fuŭ	Wu,	'U	Tsĭaŭ	Pa	Kĭung
27	Wu	Ngue	Sran	Ku,	Trhĭa	Hiŭ	Fu	Bung
28	Dzĭŭan	Trhi	Pan	Ngĭang,	Tshiŭ	Drĭung	Yi	Kĭung

Values correspond to Yuan dynasty pronunciation rather than modern pronunciation:

- r indicates the previous consonant is pronounced with a curled tongue tip.
- h indicates that the previous sound is pronounced with an extra puff of air.
- ny is as in 'onion', ng as in 'finger'.
- ' is a glottal stop, the sound in the middle of 'uh-oh'.
- x indicates a sound like the 'h' in 'huge'.
- A vowel with a ˘ mark means that it is a "glide" – a short vowel-like sound transitioning into or out of the syllable's main vowel: so ŭ and ŏ are like a 'w', while ĭ is like a 'y'.

Q4.1. In your answer book, fill in the 18 missing names. (10 points)

Q4.2. On the right is a partial 3x3 excerpt from one larger page of a 1418 manuscript of the Băijiāxìng Měnggǔwén. Six of the names have been left out. In your answer book, draw them in the spaces provided. (9 points)

Q4.3. Explain how the ʹPhags-pa writing system works. (6 points)
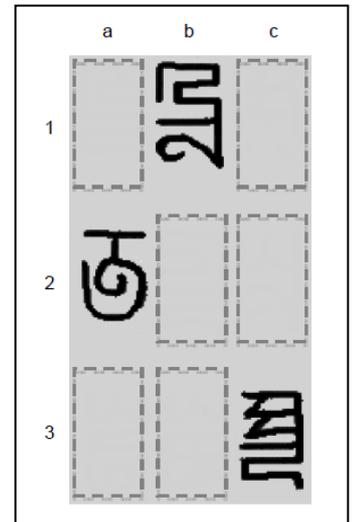
Q4. 100 surnames

4.1

13a.		15f.		19e.	
13g.		15g.		20b.	
14c.		16g.		20c.	
14e.		17a.		22c.	
15b.		18b.		23c.	
15d.		18g.		24c.	

4.2		6

4.3	

Q4. 100 surnames

4.1

13a.	Kĭi	15f.	Kĭue	19e.	Tsam	3
13g.	Maŭ	15g.	Ma	20b.	Fang	3
14c.	Srĭu	16g.	Yan	20c.	Giŭ	3
14e.	Xĭang	17a.	Mue	22c.	Tso	3
15b.	Ngĭuan	18b.	Xĭa	23d.	Hŭa	3
15d.	Min	18g.	Ling	24c.	'U	3

4.2		6
4.3	To begin, we can notice that, given the complexity of the glyphs and the frequent recurrence of sub-glyph shapes within them, that the glyphs likely consist of multiple sub-parts. We can then notice that the diversity of shapes at the top of the glyphs is much greater than the shapes at the bottom: there are a few dozen possible "tops" and only a handle of possible "bottoms". This, when compared to the names in the table, shows that the top shapes of glyphs represent the possible beginnings of names, of which there are many, and the bottoms represent the ends, of which there are few. This imbalance helps to confirm our earlier insight: that sub-glyphs represent sub-parts of names in some systematic manner. It also gives us a hypothesis about the overall writing direction of the text: that it might be top-tobottom overall as well. At this point, there are several ways to try to fix what part of the poem these two pages represent. One way is by rhyme scheme: given that the poem is made of 8-name rhyming couplets, there should be a pattern of repeating "bottom" shapes every eight characters. Given the 8x7 shape of the pages, if the text is written horizontally (in either direction), this pattern should occur in vertical lines, and if the text is written vertically (in either direction), this patterns should occur in diagonal lines.	10

It is easy to see that there are no such vertical-line patterns (especially not at the ends of lines where we would hope them to be). There is, however, an every-eight-characters diagonal rhyme pattern running through the text: from 3rd row/1st column ("3a") upward and rightward to (1c), wrapping around to 7e and then up to 1k, wrapping around again to 7m, then up to 4p.

Going by the bottom shape, the rhyme scheme of the segment pictured here looks to be something like AAAAABBBBAAAAA. There is only one place in the poem this could be: the BBBB glyphs have to represent Kŭaŭ, Laŭ, Faŭ, and Maŭ and the rest represent names ending in -ng. This result, however, is still compatible with two writing directions (upward then left-to-right, or downward then right-to-left). The former of these is unlikely given the downward writing direction within glyphs, but it is still at least possible. Looking at the "top" shapes decides it for us: we have an ABCDEFAGCHGIDJ pattern, and this only fits one way with the pattern of names beginning the text (the downward then right-to-left direction).

This is one way of determining the position and direction of the manuscript relative to the poem; other ways are equally valid and can receive equal points.

Now that we know which glyphs represent which, we can begin to determine what exactly the sub-glyph shapes represent. If you've come this far, this process will be mostly straightforward. There are four "classes" of glyphs, which I'll call A, B, C, and D:

A. initial sounds (representing word-initial b, p, dz, kh, y, tr, tsh, etc.)

B. "on-glides" (ĭ, ŭ, etc.), written after initials

C. vowels (i, u, e, etc.), written after on-glides (if any)

D. "tails"/"codas" (m, ng, ĭ, etc.), written last.

Depending on where in a word a sound occurs, a sound like [ng] or [ŭ] might have to be in a different class and thus have a different shape.

One complication to watch out for is that there is no shape for [a]. If there is no other class C (vowel) glyph in the word, that word's vowel is [a]. You can tell the difference between (say) Hŭa and Haŭ, even though [a] is not written, by the shape [ŭ] gets: the class B or class D shape.

The other complication to note is that when class B [ŭ] and [ĭ] co-occur, they are written in 'Phags-pa in an order opposite from what we would expect.