

Структурная иероглифика. Композиционный ввод тангутских иероглифов.

Путько Б.М.

Принципы структурного подхода к иероглифике были сформулированы при разработке композиционного ввода китайских иероглифов упрощённого начертания [1]. Технология композиционного ввода признана изобретением и защищена патентами РФ (РФ патент № 2671043) и США (US Patent No. 10,656,830). При этом следует иметь в виду, что патентование само по себе делает всего две вещи:

1. Свидетельствует о том, что до автора об этом, если кто и догадывался, то ничего подобного не реализовал. Т.е. фиксирует право первого связного изложения идеи изобретения.
2. Свидетельствует о том, что данное изобретение соответствует объявленной задаче и решает её. Т.е. фиксирует право первой реализации идеи изобретения.

Однако, в научных исследованиях принцип проверяемости никто не отменял, и отсюда совершенно естественно возникла задача распространить найденные закономерности и проверить их работу на других схожих системах письма. И совершенно естественно таким "испытательным полигоном" стало тангутское письмо: обилие оригинальных текстов, ожидающих публикации, необходимость массированного ввода этих оригинальных текстов, необходимость анализа и обсуждения этих текстов, и, самое главное, необходимость разработки электронных словарей тангутского языка. И какой бы качественной ни была лексикографическая основа этих потребных словарей, без эффективной системы ввода иероглифов с оригинального носителя или его копии, значение подобных словарей с примитивной системой поиска иероглифов в индексах, хоть графических, хоть фонетических, "на бумаге" было бы уничтожающе малым. Три клика мышкой для поиска любого иероглифа (как в современном китайском) это, с одной стороны, идеальный вариант, с другой стороны — вполне достижимая реальность.

Постановка задачи.

Структурная иероглифика исходит из анализа графической структуры иероглифов и полагает, что каждый иероглиф можно описать с помощью некой линейной последовательности отдельных неделимых графических единиц из наперёд заданного набора таковых, каждая из которых может быть поставлена в соответствие отдельному компоненту в составе иероглифа. Эти единицы, кодирующие компоненты иероглифов в структурной иероглифике называются *графами*. Задача, следовательно, состоит в том, чтобы выделить набор возможных компонентов, из которых состоят тангутские иероглифы, определить набор графов, с помощью которых можно закодировать эти компоненты и с помощью графов описать последовательность компонентов каждого иероглифа. Для терминологической определённости здесь и далее выбранные нами для разделения иероглифов компоненты будем называть *знаками*. То есть иероглифы состоят из знаков, знаки кодируются графами. Последовательность описания компонентов иероглифов, что в китайской иероглифике традиционно называют "правила декомпозиции", предельно просты: слева-сверху вправо-вниз, для охватывающих знаков снаружи-внутри. Таким образом структурная иероглифика разворачивает двумерную структуру иероглифа в линейную последовательность графов (𐞑: 𐞑 𐞑 𐞑 𐞑 𐞑). Понятно, что знаков может быть больше, чем количество кодирующих их графов, но в составе каждого конкретного иероглифа каждый конкретный граф кодирует один конкретный знак. Граф чаще всего визуально совпадает с тем знаком, который он кодирует, но иногда он кодирует ещё и дополнительные знаки, несколько отличающиеся от изображения основного знака, но ему подобные. Эти знаки будем называть *вариантами* основного знака. Например, граф 𐞑 кодирует основной знак 𐞑 и ещё три его варианта: 𐞑, 𐞑, 𐞑 (см. приложение).

Технически задача состоит в том, чтобы составить 2 таблицы: таблицу графов с указанием знаков и их вариантов, которые они кодируют, и таблицу иероглифов с указанием последовательности графов, описывающих каждый иероглиф. Первая таблица, таблица графов со всеми вариантами знаков нужна для нас, а вторую таблицу, таблицу иероглифов, нужно сообщить машине.

таблица 1.

Фрагмент таблицы графов. Полный вариант см. приложение.

| № п/п | граф | варианты | примеры |
|-------|------|----------|---------|
| - | 𠂇 | 𠂇 𠂇 | 𠂇 𠂇 𠂇 |
| - | 𠂈 | 𠂈 | 𠂈 𠂈 |
| - | 𠂉 | 𠂉 | 𠂉 𠂉 |

таблица 2.

Фрагмент таблицы иероглифов
(иероглифы и описывающие их последовательности графов)

| | |
|---|---------|
| 𠂇 | 𠂇 𠂇 |
| 𠂇 | 一 干 彡 |
| 𠂇 | 一 𠂇 川 |
| 𠂇 | 一 彡 匕 |
| 𠂇 | 一 川 彡 匕 |

Обе эти таблицы дают нам возможность понять, что нужно спросить у машины, чтобы найти (и ввести) тот или иной иероглиф, и представляют возможность машине правильно ответить на правильно сформулированный запрос. Мы можем спросить, например, какие иероглифы содержат в своём описании в любой позиции граф 𠂇? Машина ответит нам, что таковых 20 штук (𠂇 𠂇 𠂇 𠂇 𠂇 𠂇 ...). А в комбинации с графом 𠂇? Всего в двух иероглифах (𠂇 𠂇). А в каких иероглифах граф 𠂇 находится в последней позиции? В 17-ти иероглифах (𠂇 𠂇 𠂇 𠂇 𠂇 𠂇 ...). А в первой позиции? Всего в трёх (𠂇 𠂇 𠂇).

Способы выделения компонентов иероглифов могут казаться произвольными, ибо разделить почти каждый иероглиф, особенно сложный иероглиф, на составные части можно несколькими способами, но "произвол" выделения компонентов ограничен задачей — созданием эффективной системы ввода. Отсюда знаки в первую очередь должны быть *очевидными*, т. е. они должны легко вычлениваться визуальнo в составе иероглифа, они должны быть легко заметны в составе иероглифа. Набор знаков должен быть *необходимым* и *достаточным* для представления всех рассматриваемых иероглифов. Знаки должны легко объединяться в группы по подобию для возможности кодировать одним графом набора их вариантов. При этом количество графов, кодирующих весь набор знаков, может варьировать, но должно быть обзорным. Как показал опыт создания метода композиционного ввода для китайских иероглифов, количество графов больше двух сотен становится неоперабельным. Следовательно мы можем предварительно принять, что количество графов для тангутского письма должно быть около двухсот.

Тангутские иероглифы отличает явная простота, если не сказать бедность, изобразительных ресурсов наряду с повышенным вниманием к мелким деталям начертания, которые кажутся слишком "не жизненными" в реальном письме. Два примера: 𠂇 — "корова", 𠂇 — "ценности, имущество"¹. Зачем в иероглифе 𠂇 две точки в последнем знаке? Ничего подобного без этих точек нет, он один такой и вводится простым указанием первых двух знаков. Из всего набора иероглифов всех случаев, где встречаются эти злосчастные точки всего четыре: 𠂇, 𠂇, 𠂇, 𠂇. Понятно, что для создателя это имело какое-то значение, но нам нынче оно недоступно, и нам нужно не только не потерять эти "важные мелочи", но и не утонуть в них. Самым простым решением представляется ослабление требования уникальности описания иероглифов. Таким образом мы увеличиваем риск неопределённости при вводе некоторых иероглифов и принимаем на себя ответственность за выбор нужного (см. обсуждение далее).

1- Здесь и далее все значения приведены по словарю Кычанова [2].

Реализация.

Материалом для анализа структуры иероглифов послужили 6146 иероглифов из разделов Tangut и Tangut Supplement Юникода версии 15.1. В результате была построена таблица графов, состоящая из 205 единиц (см. приложение 1), и полная таблица 6146-ти иероглифов. И та, и другая таблицы позволили организовать простой и эффективный метод выбора иероглифов, позволяющий организовать почти любую комбинаторику графов (и, соответственно, компонентов).

Компьютерная программа для ввода тангутских иероглифов была реализована на языке Java, что позволяет использовать её в основных операционных системах Windows, Linux и MacOS. Весь ввод тангутских иероглифов в данном тексте и всё обсуждение особенностей тангутской иероглифики основано на применении данной системы ввода [3].

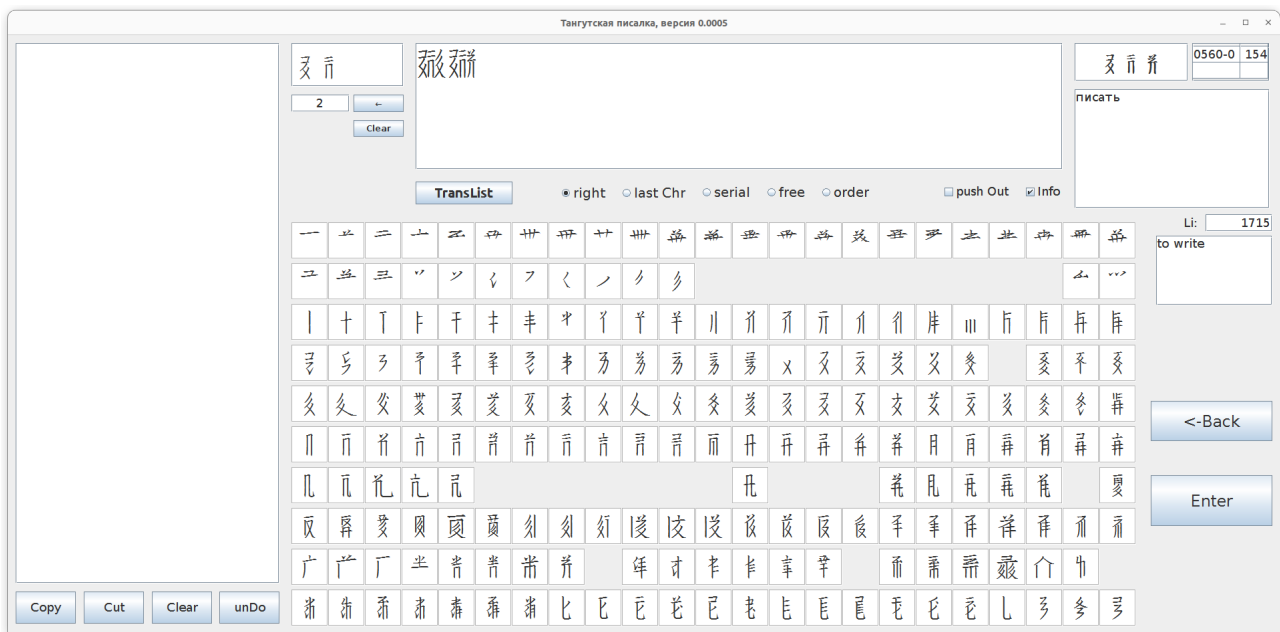


Рис.1. Интерфейс программы ввода тангутских иероглифов.

По сути, это виртуальная клавиатура со своим фокусом ввода, откуда введённый текст через буфер обмена можно переносить куда угодно. В качестве дополнительной информации по каждому иероглифу представлено его полное описание (последовательность графов) и, там, где это возможно, его номер по словарю Кычанова [2], номер страницы и перевод, а также номер иероглифа в словаре Ли Фаньвэня [4] и его перевод на английский экспортированные из программы BabelPad [5] Эндрю Веста, за что ему отдельное спасибо. Программа требует установки шрифта Tangut Yinchuan², поскольку четыре графа (𐰃 𐰄 𐰅 𐰆) расположены в области Private Use Area этого шрифта и другими шрифтами могут не отображаться.

В системе ввода используются пять режимов выборки, которые обеспечивают почти все мыслимые способы поиска компонентов в иероглифах: в начале, в конце, в середине, в комбинациях с другими знаками (см. рис. 1).

Первый режим — *right*, достался в наследство от режима *fsl* ("first-second-last") китайского метода позиционного ввода. Этот режим назван "правильным", потому что это самый эффективный режим выборки для ввода иероглифов. Вводим первый, второй и последний граф и получаем искомый иероглиф. Кроме небольшого количества случаев (см. обсуждение ниже), когда в поле иероглифов остаётся некоторая *неопределённость*, а именно, два "кандидата", и искомое нужно выбрать из двух "вручную". Вместо этого можно ввести ещё один *предпоследний* граф, и в поле иероглифов остаётся единственный уникальный иероглиф.

Второй режим — *last Chr*, последний символ, предполагает ввод графов в обратном порядке, с последнего до первого. Режим полезен для тех, кто привык работать с русской графической системой, по которой построен, например, словарь Кычанова [2].

Третий режим — *serial*, последовательный ввод графов один за другим, но не обязательно с первого. Это режим выбора по подобию, он позволяет выбирать иероглифы, имеющие общие блоки знаков.

2 - This font has been made available for free download and use with the permission of Prof. Jing Yongshi, who reserves the copyright to the font. Ссылка для скачивания: <https://www.babelstone.co.uk/Fonts/Download/TangutYinchuan.ttf>.

Четвёртый режим выбора — *free*, самый общий, он позволяет вводить графы в любом порядке. Система покажет все иероглифы, в которых присутствуют все введённые графы в любой последовательности.

Последний режим — *order*, показывает иероглифы, в которых введённые пользователем графы следуют друг за другом, но не непосредственно, а позволяют любым другим знакам встраиваться между ними — по порядку, но не обязательно один сразу за другим.

Учитывая предварительный характер этого решения, стоит всё же сказать, что это достаточно эффективный инструмент для изучения структуры тангутских иероглифов, ибо оно позволяет рассматривать реальную графическую структуру иероглифов без привлечения внешних относительно графики параметров разделения иероглифов на идеограммы, фоноидеограммы и т. п. "классы", которые имеют исключительно мнемонический смысл и упрощают запоминание иероглифов и их значений, но не имеют никакого отношения к реальной графической структуре иероглифа.

Тангутское письмо.

Уже первый предварительный взгляд на массив тангутского текста позволяет увидеть его особенности по сравнению с китайской иероглификой. Этакая смесь примитивности и изошрённости. Простым набором вертикальных, горизонтальных и наклонных черт создаётся впечатление чего-то сверхсложного и непонятного. Совершенно отсутствуют замкнутые элементы. Кажется, что пишущий старательно избегает необходимости вернуться кисточкой туда, откуда он начал. Боится не попасть? Да нет, скорее боится попасть. Замкнутые контуры — характерный приём китайского письма. (Самый частый знак в китайских иероглифах — "рот". А если добавить к нему "ограду", "глаз", "солнце", "белый", "поле"... в 77-и графах из 254-х, описывающих 20992 иероглифа раздела CJK Unified Ideographs Юникода, присутствуют замкнутые контуры.) Кажется, основная задача создателей тангутских иероглифов заключалась в том, чтобы результат ни на йоту не походил на китайские иероглифы.

Совершенно непритязательны способы создания новых знаков. Есть один иероглиф 𠂇 ("один"), развернём его знаки в обратном порядке и сделаем из него новый — 𠂈 ("делать"). Есть иероглиф 𠂉 ("копить"), накроем его крышкой, тоже получим нечто новое — 𠂊 ("вянуть"). Берём иероглиф 𠂋, меняем средний компонент, получаем иероглиф 𠂌. Ещё одна замена, получаем 𠂍.

𠂋 - наказание; штраф.

𠂌 - стареть; портиться; разрушаться.

𠂍 - поносить; позорить; злословить.

При отсутствии длительной, многовековой традиции видеть разнообразный и богатый набор знаков (достаточно только упомянуть 540 ключей словаря Сюй Шеня), и при таком изначально простом наборе изобразительных ресурсов (всего 14 иероглифов, состоящих из одного компонента, против 315 в китайском), требует активного использования комбинаторики. И всё это "вручную", по памяти. Из 6146 иероглифов 1392 состоят из двух компонентов. Среди двухкомпонентных тангутских иероглифов 296 иероглифов-палиндромов (перевёртыша): 𠂎 - 𠂏, 𠂐 - 𠂑, 𠂒 - 𠂓, 𠂔 - 𠂕 и т. д.. Каждый пятый — палиндром. И это, как вы понимаете, не всё. Есть ещё иероглифы из трёх и более компонентов, а среди них тоже присутствуют палиндромы (𠂖 - 𠂗, 𠂘 - 𠂙, 𠂚 - 𠂛 и т.д.)³.

В результате такой немудрёной стратегии создания тангутские иероглифы явным образом состоят из одной, а чаще всего из двух и трёх колонок знаков. Некоторые наиболее изошрённые содержат до пяти колонок. Колонки эти состоят каждая из одного, двух, редко трёх знаков. В наборе знаков заметно выделяются те, которые могут "расползаться" на соседние колонки (𠂜 𠂝 𠂞). Эти "расползающиеся шляпки" в рукописном тексте — кошмар для читателя. 𠂟 и 𠂠 - разнописи одного и того же иероглифа ("кошмар; видеть во сне кошмары; удушье во время сна; сон; иллюзия"). А эти два: 𠂡 и 𠂢, - разные. Первый — "(ночной) патруль", а второй — "тёмная, глухая ночь".

Оказавшись внизу иероглифа, эти знаки могут оставаться в рамках своей колонки, а могут подчёркивать все колонки иероглифа (𠂣 𠂤). При этом, нижние подчёркивающие знаки часто сами подчёркиваются дополнительной горизонталью. Знак 𠂥 внизу иероглифа подчёркнут всегда (𠂦 𠂧). Знак 𠂨 внизу подчёркнут почти в 2 раза чаще (45:27), чем оставлен не подчёркнутым (𠂩 𠂪). Знак 𠂫 внизу подчёркнут

3 - Для сравнения, из 20992 китайских иероглифов (CJK Unified Ideographs) 3641 иероглиф состоит из двух компонентов. Среди них 40 палиндромов, и из этих сорока 16 являются простой заменой друг друга, 12 отсутствуют в современных словарях (это значит, что в реальности они никому не нужны). После двух тысяч лет употребления в иероглифике осталось только 12 иероглифов-палиндромов против 296 в тангутском.

всегда (𠃉). Но все эти подчёркивания никакого смыслообразующего значения не имеют, поэтому их можно смело отнести к вариантам основных знаков без подчёркиваний (см. приложение). Также следует отметить, что такой растянутый знак внизу необходимо рассматривать как последний знак последней колонки.

Явным образом выделяется группа п-образных знаков со сходным типом модификации: 𠃉 и 𠃊, 𠃋 и 𠃌, 𠃍 и 𠃎, 𠃏 и 𠃐, 𠃑 и 𠃒, и т. д. Знаки с нижним крюком вправо очень часто, но не всегда, расположены в нижней правой части иероглифа (𠃉 𠃊 𠃋 𠃌 𠃍 𠃎). Как видно на приведённых примерах, такой знак с крюком внизу часто сопровождается дополнительной откидной справа. Но из 165-ти знаков с откидной только в шести случаях, когда знак 𠃉 приобретает вид 𠃉, наличие этой откидной обретает смысл: 𠃉 - трудный, 𠃉 - спрашивать. Далее будем считать знаки с откидной вариантами знаков без таковой (см. приложение 1).

Подобны по способу образования вариантов и знаки 𠃉 и 𠃉. Первый из них встречается в различных иероглифах 95 раз и только дважды с нижним крюком вправо в иероглифах 𠃉 и 𠃉. Второй знак встречается в иероглифах 140 раз. По аналогии с п-образными знаками в правом нижнем положении 66 раз он имеет нижний крюк (𠃉), а 20 раз в этом же положении свободно обходится без него (𠃉).

Существует отдельный набор знаков, которые никогда не появляются в первой позиции. Это знаки: 𠃉, 𠃉, 𠃉, 𠃉, 𠃉, 𠃉, 𠃉. Почему, бог весть.

Наиболее интересной представляется группа завершающих знаков. Совершенно однозначно можно утверждать, что в китайской иероглифике таких знаков нет. Сводка завершающих знаков представлена в таблице 3. Знаки указаны в том же порядке, как на матрице графов (см. рис. 1). Знаков этой категории 17, причём последний знак 𠃉 в системе описывается как последовательность двух графов: 𠃉, 𠃉. Он явно отличается от знака в 14-ом ряду таблицы. Обратите на это внимание. В колонке "в иной позиции" указано количество соответствующего знака в незавершающей позиции иероглифа, в колонке "исключения" представлен исчерпывающий список иероглифов, в которых соответствующий знак не является завершающим по формальным основаниям, а именно, добавлена одна или две горизонтали "поверх" его. Но визуально соответствующий знак во всех примерах (за исключением трёх случаев с нижним знаком 𠃉) занимает правый нижний угол иероглифов и спокойно может рассматриваться как завершающий. Это и даёт нам право относить эту группу знаков (хоть и с некоторой условностью) к категории завершающих.

таблица 3.

Завершающие знаки.

| знак | в последней позиции | всего | в иной позиции | исключения |
|------|---------------------|-------|----------------|-----------------|
| 𠃉 | 584 | 584 | - | |
| 𠃊 | 241 | 241 | - | |
| 𠃋 | 84 | 84 | - | |
| 𠃌 | 69 | 78 | 9 | 𠃌 𠃌 𠃌 𠃌 𠃌 𠃌 𠃌 𠃌 |
| 𠃍 | 48 | 49 | 1 | 𠃍 |
| 𠃎 | 73 | 78 | 5 | 𠃎 𠃎 𠃎 𠃎 𠃎 |
| 𠃏 | 36 | 37 | 1 | 𠃏 |
| 𠃐 | 35 | 36 | 1 | 𠃐 |
| 𠃑 | 47 | 54 | 7 | 𠃑 𠃑 𠃑 𠃑 𠃑 𠃑 𠃑 |
| 𠃒 | 20 | 20 | - | |
| 𠃓 | 25 | 26 | 1 | 𠃓 |
| 𠃔 | 10 | 10 | - | |
| 𠃕 | 15 | 22 | 7 | 𠃕 𠃕 𠃕 𠃕 𠃕 𠃕 𠃕 |

| знак | в последней позиции | всего | в иной позиции | исключения |
|------|---------------------|-------|----------------|------------|
| 𠄎 | 66 | 66 | - | |
| 𠄏 | 66 | 66 | - | |
| 𠄐 | 72 | 73 | 1 | 𠄑 |
| 𠄒 | 13 | 14 | 1 | 𠄓 |

Посмотрите на структуру следующих иероглифов: 𠄎, 𠄏, 𠄐, 𠄑, 𠄒. Общий знак 𠄎, и в дополнение к нему четыре разных знака. Непроизвольно ожидается некая семантическая база с различными нюансами. Но нет.

𠄎 - белый нефрит.

𠄏 - корень.

𠄐 - делить; серп.

𠄑 - различаться; указать; понять; обнаружиться.

𠄒 - я; служебное слово (суффикс).

Конечно, стоит оговориться, что в третьем примере знак 𠄐 не относится к категории завершающих, но пока я не о том, общая структурная схема налицо. И полное отсутствие семантической связи. Это и обращает на себя внимание в завершающих знаках. Складывается впечатление, что основная задача завершающего знака сказать, "нет, это не то, что вы подумали". Но и это впечатление обманчиво. В большинстве случаев завершающий знак не имеет смысловозначительного значения вообще. Из 584 случаев употребления самого распространённого завершающего знака 𠄎 только 102 отличают один иероглиф от другого. В остальных 482-х случаях этот знак не нужен, это метка неизвестно чего, кроме желания их создателя. Это может быть влияние эстетических, религиозных, каких угодно иных представлений, но никак не желание использовать его как признак отличия для двух (или нескольких) семантических полей.

Знаки 𠄎 и 𠄏. Второй кодируется двумя графами 𠄎, но оба знака завершающие, нигде не пересекаются и не конкурируют друг с другом при вводе. Можно было бы спокойно объявить их вариантами одного знака, но пока оставим как есть. Явное визуальное различие между этими хоть и похожими знаками может обескураживать не подготовленного пользователя. С регулярным завершением рукописного знака ситуация не уникальна. Отмечаются подобные варианты знаков 𠄑 и 𠄒, 𠄓 и 𠄔, когда количество горизонтальных чёрточек варьирует либо по ошибке, либо по невнимательности. А поскольку на значение иероглифа это никак не влияет, то и может казаться пишущему не важным при воспроизведении.

Всего завершающие знаки используются в 1504 иероглифах. Практически каждый четвёртый иероглиф заканчивается таким знаком.

Обсуждение.

То, что все иероглифы состоят из знаков, ни у кого не вызывает сомнения ("說文解字"). Сомнения начинаются тогда, когда пытаются выяснить, из каких же знаков состоят иероглифы. "Но тангутские филологи, которые в отличие от Сюй Шеня не имели возможности указать исходных простых знаков в естественном процессе развития письма, объясняли сложные знаки через сложные же. Тангутскими филологами был использован именно принцип членения знака на составные элементы, впервые предложенный Сюй Шенем." [6, с. 16].

Евгений Иванович Кычанов в "Очерке истории тангутского государства" 1968-го года [7] упоминал "идеографообразующие элементы тангутского письма", а в "Истории тангутского государства" [8, с. 408] привёл список из 83-х "идеографообразующих элементов". Но ввод конкретных иероглифов с таким именно набором "идеографообразующих элементов" был бы неоправданно долгим, а с учётом анализа взаимного расположения этих элементов по двадцати типам превращался бы в утомительную, хотя и оригинальную головоломку. Задание по примеру китайской иероглифики направления декомпозиции иероглифов "сверху-слева вправо-вниз" сразу кардинально упростило ситуацию. Двадцать типов взаимного расположения знаков в иероглифах выродились в линейную последовательность. А возможность спросить у машины вхождение каждого конкретного знака в любой позиции сделало решение задачи поиска (ввода) иероглифа тривиальной. Даже простое последовательное перечисление знаков, входящих в иероглиф, гораздо эффективнее их прорисовки "по чёрточкам", ибо указание целых "блоков чёрточек" даже методом проб и ошибок гораздо эффективнее метода рукописного ввода, не говоря уж о качестве "прорисовки" иероглифов. Конечно, алгоритмы выборки, отработанные на китайских иероглифах не обязаны быть столь же эффективными для иероглифов тангутских.

Эффективность режима выборки *fsl* в нынешней китайской иероглифике — три клика на иероглиф, за исключением 112-ти иероглифов, когда при вводе требуется четвёртый клик. Другими словами, существует 112 иероглифов, которые однозначно не определяются по первому, второму и последнему графам. Ввод предпоследнего графа частично решает эту проблему для 102-х иероглифов. И только 10 иероглифов из 6724-х оставляют неопределённость после ввода предпоследнего графа и оставляют решение пользователю. В китайском упрощённом эта проблема решена с помощью т. н. *маркёров*, уникальных, характерных для каждого иероглифа последовательностей не больше трёх графов (см. [1], с. 24). Использование маркёров позволяет все китайские иероглифы вводить максимально за три клика. Для тангутских иероглифов маркёры пока не выработаны, и неопределённости в режиме *fsl* заметно больше, а именно, ввод четвёртого графа, предпоследнего, требует 254 иероглифа, в два с небольшим раза больше, чем в китайском упрощённого начертания. И связано это, во-первых, с наличием в тангутских иероглифах "завершающих знаков", у которых смысловозначительная нагрузка по воле создателей значительно снижена, а в контуре иероглифа они присутствуют (𐰇 𐰈 𐰉 𐰊 𐰋 𐰌 - и т. д.). Во-вторых, это связано с общей тенденцией к стремлению создателя иероглифов прятать различие в иероглифах внутрь. Количество тангутских иероглифов в блоках с подобным контуром, в котором различия спрятаны, может достигать до трёх десятков (𐰍 𐰎 𐰏 𐰐 𐰑 𐰒 𐰓 𐰔 𐰕 𐰖 𐰗 𐰘 𐰙 𐰚 𐰛 𐰜 𐰝 𐰞 𐰟 𐰠 𐰡 𐰢 𐰣 𐰤 𐰥 𐰦 𐰧 𐰨 𐰩 𐰪 𐰫 𐰬 𐰭 𐰮 𐰯 𐰰 𐰱 𐰲 𐰳 𐰴 𐰵 𐰶 𐰷 𐰸 𐰹 𐰺 𐰻 𐰼 𐰽 𐰾 𐰿 - и т. д.). В китайском тоже есть подобные блоки иероглифов (唼 唺 唻, 羸 羸 羸 羸), но их нужно ещё поискать, а в тангутском они на каждом шагу.

Казалось бы, какая разница, как компоновать новый иероглиф? Дело в том, что распознавание сложных графических объектов в значительной степени связано с анализом контуров этих объектов. Зрительная кора головного мозга эволюционно устроена таким образом, что контуры объектов анализируются раньше их наполнения, сразу после выделения их из фона [9]. Объекты разнящиеся контурами распознаются быстрее и эффективнее, с меньшими затратами "вычислительных ресурсов" мозга. Отсюда и расположение ключей китайских иероглифов слева, сверху, справа, внизу и редко в центре. Отсюда и внимание русской графической системы на последнюю черту иероглифа, отсюда и кодировка иероглифов по четырём углам. И, в общем случае, свидетельство Василия Михайловича Алексева: "Всякий, кто видел особым жирным почерком написанные китайские вывески, согласится, что он "читал" в сущности по контурам знаков, не замечая этого" [10, с. 131, прим. 75a]. Китайская иероглифика "явочным порядком" пришла к пониманию этого принципа. Внимание в процессе зрительного восприятия объекта при отсутствии различий по контуру вынуждено переходить внутрь знаков и искать различия там. Этот подход к анализу зрительных образов и реализует алгоритм *fsl*: два первых и два последних графа, остальное не важно. В тангутском письме, увы, это не так. Точнее, бывает и так, но значительно реже, чем в китайском. Многие поколения китайских писцов не только массово писали иероглифы, но и массово забывали то, что плохо воспринималось и то, что легко забывалось. И по необходимости заменяли одни иероглифы на другие или модифицировали их по своим собственным соображениям. Часть этих изменений воспринималась другими писцами и сама становилась "стандартом". "Археологические материалы свидетельствуют о том, что стандартизация китайской письменности была постепенным и не всегда однонаправленным процессом, длившимся несколько столетий" [11, с. 144]. Тангутское письмо было лишено этого длительного периода "естественного отбора" памятью пишущих и обучаемых.

Опыт реального ввода гнездовых иероглифов словаря Кычанова [2] в качестве справочного материала в данной системе показал, что в корпусе из 5797 присутствует 86 иероглифов, для которых зафиксировано наличие разнописей. В реальных рукописных текстах их, конечно же, больше. При этом совершенно избежать ошибок переписчика тоже не могли. Следовательно, при разработке системы ввода тангутских иероглифов, а равно любого другого мёртвого языка, в отличие от живого китайского, необходимо сместить некоторые акценты. И это важно настолько, что незначительные методические изменения становятся изменениями методологическими. Поскольку тангутский язык давно мёртв, в публикациях в первую очередь требуется точность передачи текстов, зафиксированных на носителе. В мёртвых текстах уже нет никаких "правил орфографии", даже если их создатели когда-то и декларировали таковые. И теперь для нас существует одно лишь правило — "правильно написано так, как написано в документе". Чтобы было проще воспроизвести особенности документа: все "неправильные" разнописи, все "ошибки" автора или переписчика, - и при этом не плодить мелкие особенности изображений знаков, проще всего ослабить требование уникальности описания иероглифов в таблице иероглифов. В этом случае два (очень редко три) иероглифа, имеющие совпадающие последовательности графов в их описаниях, при вводе будут представлены на выбор пользователя. Это снижает скорость ввода, но "обостряет внимательность", что весьма важно, ибо точность ввода критически важнее его скорости. И, вводя иероглиф неясного начертания 𐰿, полезно видеть, что существует ещё один очень на него похожий 𐰿, и выбрать в данном случае подходящий. Таких пар иероглифов с совпадающими описаниями 94 шт. Хотя вариантов написания знаков гораздо больше, и все они являются важным материалом для дальнейшего исследования.

Понятно, что сохранность текстов зачастую оставляет желать лучшего, и знание структурных особенностей знаков может в некоторых случаях помочь восстановить утраченные части иероглифов или не обращать внимание на потерю некоторых элементов их. Например, последовательность знаков 𐰿 и 𐰿 во всём корпусе иероглифов встречается только один раз, причём в первой позиции в иероглифе 𐰿. 24 знака никогда не встре-

чаются в первой позиции иероглифа. Из 17-ти завершающих знаков семь никогда не встречаются в иной позиции. Знание такого рода простой комбинаторики знаков может реально помочь исследователям в восстановлении утраченных участков текста. Возможность подключения к системе ввода сопутствующей информации по каждому иероглифу (например, перевода его по словарю Кычанова [2]) позволяет тут же "на лету" проверять обоснованность собственных предположений.

Практика применения подобной системы ввода не только говорит о возможности подключения композиционного ввода к электронным словарям, но и напрямую демонстрирует его преимущество перед "традиционными" методами поиска в словарях "по ключам", "темам" и разного рода указателям. Достаточно ввести иероглиф в поле поиска, его значение в словаре мы получаем практически мгновенно.

Предварительный анализ состава иероглифов выявляет значительное количество среди иероглифов перевертышей, бессистемное и непонятное употребление мелких деталей знаков то как чисто стилистических, то как несущих строгое отличительное значение (𠄎 и 𠄏). Примеров тому приведено достаточно: п-образные знаки в последней позиции иероглифа, последняя откидная у таких знаков с крючком вправо, произвольное и бессистемное подчёркивание "гребёночково-образных" знаков внизу (𠄎 𠄏). Кажется совсем не очевидным, что "в процессе сложения письма каждому идеографообразующему элементу присваивалось строго определённое смысловое значение" [7, с. 265]. Наверное, это было так. Точнее, это могло быть и так. А потом спокойно менялось создателем (или переписчиками) по мере надобности на всё, что угодно.

Какие "строгие размышления" о смысловом наполнении или фонетической передаче могли привести создателя тангутской письменности к использованию знака 𠄎 в 1286 иероглифах, в каждом пятом? Какие реалии тангутской жизни или тангутской фонетики может отражать такое количество перевертышей? Детерминатив в них превращается в фонетик, а фонетик в детерминатив? А потом ещё переписчики начинают путаться с такого рода комбинаторикой: 𠄎 — красивый, пёстрый, цветущий и 𠄏 тоже красивый, пёстрый, цветущий.

Как бы то ни было, теперь мы имеем механизм простой проверки такого рода предположений и отсеивания из анализа явных заблуждений. А с возможным теперь появлением электронных словарей тангутского языка проблема анализа сложной и изощрённой мнемотехники "Моря письмен" совершенно теряет актуальность.

Возможности (и ограничения) механизмов зрительного восприятия сложных графических объектов породили феномен иероглифики как исторически первого и самого длительно существующего способа передачи смысла. А применение компьютера даже человеком неподготовленным, как автор данного текста, не знающим ни звучания, ни значения иероглифов, позволяет вводить иероглифы, ориентируясь только на их очевидную графическую структуру.

Выводы.

1. Принципы композиционного ввода, выработанные для китайского языка, явно продемонстрировали действенность и работоспособность и на примере тангутской иероглифики.
2. Предложен первый вполне работоспособный прототип виртуальной клавиатуры для ввода тангутских текстов.
3. Справочная система представленной виртуальной клавиатуры наглядно демонстрирует возможность в дальнейшем создания и использования электронных словарей тангутского языка.

Список литературы.

1. Путько Б.М., "Введение в структурную иероглифику". – Lambert Academic Publishing, 2018. Последняя исправленная и несколько дополненная версия доступна по адресу <http://poutko.ru/StructChar.pdf>.
2. Кычанов Е.И. - "Словарь тангутского (си ся) языка. Тангутско-русско-английско-китайский словарь.", Kioto, Japan, 2006.
3. Путько Б.М. - ScribblerII (<https://poutko.ru/ScribblerII.zip>)
4. Ли Фаньвэнь - Ся Хань Цзыдянь, Пекин, 2008.
5. West Andrew - BabelPad (<http://www.babelstone.co.uk/Software/BabelPad.html>).
6. Море письмен. Факсимиле тангутских ксилографов. Часть 1. Москва, 1969.
7. Кычанов Е.И. - Очерк истории тангутского государства, М., Наука, 1968.
8. Кычанов Е.И. - История тангутского государства. Санкт-Петербург, 2008.
9. Й. Дудель, Й. Рэгг, Р. Шмидт и др., Физиология Человека в 3-х томах, Т.1., М., Мир, 2005.

10. Алексеев В.М., Китайская иероглифическая письменность и её латинизация. – Ленинград: Изд-во Академии наук СССР, 1932 г. (Копия доступна по адресу <https://poutko.ru/Alekseev.pdf>.)
11. Galambos, Imre. - Orthography of early chinese writing: Evidence from newly excavated manuscripts Budapest, 2006. (<http://shahon.org/wp-content/uploads/2010/02/galambos-2006-orthography-of-early-chinese-writing.pdf>)

Таблица графов.

| № п/п | граф | варианты | описание | примеры |
|-------|------|----------|---|---------|
| 1 | 一 | | Отличается только длиной, как все верхние элементы может занимать всю ширину знака. | 爾 彳 |
| 2 | 乚 | | Тоже, что предыдущий компонент. | 夂 彳 |
| 3 | 二 | | Любые 2 горизонтали. Часто для обозначения трёх горизонталей комбинируется с № 1. в порядке 1-3, как в последнем примере. | 彳 彳 彳 |
| 4 | 一 | | Чаще всего верхний элемент, соответственно может растягиваться до предела знака. | 彳 |
| 5 | 二 | | Иногда теряет маленькую перекладинку. | 彳 彳 彳 |
| 6 | 一 | 一 | Когда знак располагается внизу, подчёркивается горизонталью (№ 1). Забудьте о ней, вводить её никогда не нужно. | 彳 彳 彳 |
| 7 | 二 | 二 | Любит расплзаться в ширину знака. Если снизу подчёркнут горизонталью, как предыдущий знак (№ 6), забудьте о ней. | 彳 彳 彳 |
| 8 | 二 | | | 彳 彳 |
| 9 | 一 | | Не совсем обычный вариант во втором примере, когда вертикальные штрихи удлинены и слегка развёрнуты наружу. | 彳 彳 |
| 10 | 二 | | | 彳 彳 |
| 11 | 二 | | Эта "шляпка" однажды встречается и внизу, пример 2. | 彳 彳 |
| 12 | 二 | | Из 44 иероглифов в 31-ом комбинируется с № 90, причём в 22-х случаях сливается с ним (пример второй). | 彳 彳 |
| 13 | 二 | | Сравни со следующим (№14). | 彳 |
| 14 | 一 | | При вводе легко спутать с предыдущим знаком, если следующий за ним имеет собственную горизонталь как во втором примере. | 彳 彳 |
| 15 | 二 | | | 彳 |
| 16 | 二 | | Приблизительно поровну встречается 2 варианта: с небольшими штрихами на нижних откидных и без таковых (25:24). В примерах два иероглифа, из которых придётся выбирать вручную. При вводе остальных неопределённости не возникает. | 彳 彳 |
| 17 | 二 | | | 彳 彳 |
| 18 | 二 | | | 彳 |
| 19 | 二 | | | 彳 |
| 20 | 二 | | | 彳 |
| 21 | 二 | | | 彳 |
| 22 | 二 | | Внизу всегда подчёркнут. Эта последняя горизонталь не вводится. | 彳 彳 |
| 23 | 二 | | Тоже странная вещица. Вообще со "шляпками" нужно разбираться отдельно. Часто кажется, что это неаккуратность переписчика. | 彳 |
| 24 | 二 | 二 | Варианты разнятся только длиной нижней горизонтали. | 彳 彳 |
| 25 | 二 | | Тоже и для этого графа. См. пример 2. | 彳 彳 |
| 26 | 二 | | | 彳 |
| 27 | 二 | | Может располагаться сверху, а может охватывать знак № 38 (второй пример). | 彳 彳 |

| № п/п | граф | варианты | описание | примеры |
|-------|------|----------|--|---------|
| 28 | ノ | | | 𠂇 𠂈 |
| 29 | 彳 | | | 𠂉 𠂊 |
| 30 | フ | 𠂋 | Две основные ипостаси: вытянут донизу или слегка приплюснут. | 𠂌 𠂍 |
| 31 | く | | Очевидное использование знака в первом примере, и перечёркнутый двумя откидными вариант его же во втором. | 𠂎 𠂏 |
| 32 | ノ | | В примерах все возможные варианты знака. | 𠂐 𠂑 𠂒 𠂓 |
| 33 | 彳 | | | 𠂔 𠂕 |
| 34 | 彳 | | | 𠂖 𠂗 |
| 35 | 彳 | | | 𠂘 |
| 36 | 彳 | | | 𠂙 |
| 37 | 丨 | | В первой позиции может быть слегка изогнутой. | 𠂚 𠂛 𠂜 |
| 38 | 十 | | В вариантах знак различается только высотой. Компонуется часто с № 27. Средний знак последнего примера сборный: 十 一 二 | 𠂝 𠂞 𠂟 |
| 39 | 丨 | | Верхний горизонтальный штрих немного варьирует, иногда смещаясь в одну сторону. | 𠂠 𠂡 |
| 40 | 丨 | | | |
| 41 | 丨 | | | 𠂢 |
| 42 | 丨 | | | 𠂣 |
| 43 | 丰 | | Во втором варианте этого знака две горизонтали ушли в сторону. | 𠂤 𠂥 |
| 44 | 丰 | | Второй пример — к этому знаку добавлена откидная (№ 32). | 𠂦 𠂧 𠂨 |
| 45 | 彳 | | Вертикаль этого знака слегка "плавает" относительно откидной | 𠂩 𠂪 |
| 46 | 彳 | | | |
| 47 | 彳 | | | 𠂬 𠂭 |
| 48 | 川 | | Две вертикальных во втором примере могут быть неубедительными, но это так. В третьем примере две вертикали перечёркнуты одной короткой горизонталью. | 𠂮 𠂯 𠂰 𠂱 |
| 49 | 彳 | | | 𠂲 |
| 50 | 彳 | | Во втором примере две горизонтали (№ 3) перечёркнули две вертикальных этого знака. | 𠂳 𠂴 |
| 51 | 彳 | | | 𠂵 |
| 52 | 彳 | | То же, что и в знаке № 50. | 𠂶 𠂷 |
| 53 | 彳 | 彳 | Правая вертикаль может уйти вверх, а может и остаться под откидными (пример 2). В третьем примере 29-й знак слегка внедрился внутрь этого знака.. | 𠂸 𠂹 𠂺 |
| 54 | 彳 | | | |
| 55 | 彳 | | Слегка может обескуражить третий пример, но всё это три вертикальных. | 𠂻 𠂼 𠂽 |
| 56 | 彳 | 𠂿 𠃀 | Один вариант как во втором примере с короткой вертикалью вверх, второй вариант с отогнутой левой ногой в последнем примере | 𠂾 𠂿 𠃀 |
| 57 | 彳 | | | 𠃁 |
| 58 | 彳 | | К этому знаку добавить две горизонтали (№ 3), получим левый знак иероглифа из примера. | 𠃂 |

| № п/п | граф | варианты | описание | примеры |
|-------|------|----------|---|---------|
| 59 | 𠄎 | 𠄎 𠄎 | В примерах все варианты этого знака. Часто в последней позиции отгибает ножку как в примере 3. Знаки 57 и 58 никогда этого не делают. | 𠄎 𠄎 𠄎 𠄎 |
| 60 | 𠄎 | | | 𠄎 |
| 61 | 𠄎 | | | 𠄎 |
| 62 | 𠄎 | | Во втором примере два знака 𠄎, не путать со знаком № 203. | 𠄎 𠄎 |
| 63 | 𠄎 | | | 𠄎 |
| 64 | 𠄎 | | | 𠄎 |
| 65 | 𠄎 | | | 𠄎 |
| 66 | 𠄎 | | | 𠄎 |
| 67 | 𠄎 | | Внимательно. Второй пример может смутить, но это тот же знак. | 𠄎 𠄎 |
| 68 | 𠄎 | | | 𠄎 𠄎 |
| 69 | 𠄎 | 𠄎 | Этот знак в десяти случаях из 115-ти в последней позиции приобретает дополнительную откидную, которая имеет значоразличительный смысл лишь в одном случае (см. примеры) | 𠄎 𠄎 |
| 70 | 𠄎 | 𠄎 | Тоже относится и к этому знаку (см. пред. комм.), только соотношение иное — 15 из 108-ми, и смысла в нём никакого нет. | 𠄎 𠄎 |
| 71 | 𠄎 | | | 𠄎 |
| 72 | 𠄎 | | Верхний элемент в 21-ом случае из 30-ти расплзается вправо, имитируя шляпку. | 𠄎 𠄎 |
| 73 | 𠄎 | | Во втором примере этот знак компонуется с № 34. | 𠄎 𠄎 |
| 74 | 𠄎 | | | 𠄎 |
| 75 | 𠄎 | | | 𠄎 |
| 76 | 𠄎 | | | 𠄎 |
| 77 | 𠄎 | | | 𠄎 |
| 78 | 𠄎 | | | 𠄎 𠄎 |
| 79 | 𠄎 | | | 𠄎 |
| 80 | 𠄎 | | | 𠄎 |
| 81 | 𠄎 | | | 𠄎 |
| 82 | 𠄎 | 𠄎 𠄎 | Редкий вариант в третьем примере — две дополнительные точки. Иногда слегка укорачивается первая откидная. | 𠄎 𠄎 𠄎 𠄎 |
| 83 | 𠄎 | | | 𠄎 𠄎 |
| 84 | 𠄎 | 𠄎 | | 𠄎 𠄎 |
| 85 | 𠄎 | | Если справа есть куда ползти, гребёнка обязательно уползёт вправо. | 𠄎 |
| 86 | 𠄎 | | Разница вариантов только в длине горизонтали. | 𠄎 𠄎 |
| 87 | 𠄎 | | То же сползание верхнего элемента вправо и слегка обрезана сверху первая откидная во втором примере. | 𠄎 𠄎 |
| 88 | 𠄎 | | Здесь первая откидная обрезана всегда. | 𠄎 𠄎 |
| 89 | 𠄎 | | | 𠄎 𠄎 |
| 90 | 𠄎 | 𠄎 | В примерах два явно разных начертания этого знака, но они нигде не вызывают неопределённости. | 𠄎 𠄎 |
| 91 | 𠄎 | 𠄎 | | 𠄎 𠄎 |
| 92 | 𠄎 | 𠄎 | Варианты в количестве точек. | 𠄎 𠄎 |

| № п/п | граф | варианты | описание | примеры |
|-------|------|----------|--|-----------|
| 93 | 彡 | 彡 | Есть варианты стыковки в этом знаке верхней и нижней частей. | 彡 彡 彡 |
| 94 | 彡 | | | 彡 |
| 95 | 彡 | | | 彡 |
| 96 | 彡 | | | 彡 彡 |
| 97 | 彡 | | Дважды расширяет горизонтальную на соседнюю колонку. Чаще обходится без этого. | 彡 彡 彡 |
| 98 | 彡 | | При вводе 彡 в комбинации с 彡, шляпка не ползёт вправо, если это нужно вводите 彡 и 彡. | 彡 彡 彡 彡 彡 |
| 99 | 彡 | | | 彡 |
| 100 | 彡 | | | 彡 |
| 101 | 彡 | | | 彡 |
| 102 | 彡 | | | 彡 |
| 103 | 彡 | | В примере не совсем очевидно расположение знаков (彡 彡 彡). Привыкайте. | 彡 |
| 104 | 彡 | | | 彡 |
| 105 | 彡 | | Узнаётся всегда, даже в первом примере, слегка "нестандартное" написание во втором и вложение двух откидных внутрь в третьем. | 彡 彡 彡 |
| 106 | 彡 | 彡 | Начались "откинутые ноги" — характерный приём для правого нижнего компонента иероглифа. Иногда ногу "откидывает" элемент в середине и даже в начале иероглифа, но значительно реже. Часто такой компонент присоединяет к себе ещё и откидную, но она никогда не вводится. Забудьте об этой откидной. | 彡 彡 彡 |
| 107 | 彡 | | Часто является охватывающим знаком и впускает внутрь другие знаки. | 彡 彡 |
| 108 | 彡 | 彡 | "Откинутая нога" предыдущего знака. Чаще всего сопровождается откидной. Её также вводить не надо. | 彡 彡 |
| 109 | 彡 | | В примере две откидных внутри. Больше ничего необычного нет. | 彡 |
| 110 | 彡 | 彡 | | 彡 彡 |
| 111 | 彡 | | Внутри этого знака не бывает ничего. | 彡 |
| 112 | 彡 | 彡 | Чемпион по бессмысленным откидным - из 45-ти иероглифов всего 5 шт. без откидной. | 彡 彡 |
| 113 | 彡 | 彡 | Этот часто позволяет разместиться ещё чему-нибудь в своём чреве. Комбинация с № 38 (彡 彡) делает его плохо узнаваемым (см. пример4). | 彡 彡 彡 彡 彡 |
| 114 | 彡 | 彡 | Этот всего 4 раза из 49-ти прицепляет откидную как во втором примере. | 彡 彡 |
| 115 | 彡 | | | 彡 |
| 116 | 彡 | | Этот вообще никогда не оттопыривает ногу, даже в конце иероглифа. | 彡 |
| 117 | 彡 | | Этот знак с откинутой ногой имитируется набором "彡 彡". Да и всего-то таких случаев только два. Оба в примерах. | 彡 彡 彡 |
| 118 | 彡 | | Этот знак с откинутой ногой имитируется набором "彡 彡". И таких случаев тоже только два. | 彡 彡 彡 |
| 119 | 彡 | | | 彡 |
| 120 | 彡 | | | 彡 |
| 121 | 彡 | | | 彡 |

| № п/п | граф | варианты | описание | примеры |
|-------|------|----------|--|-----------|
| 122 | 卩 | | Любопытно, что этот знак никогда не встречается в первой позиции. | 𠂇 |
| 123 | 卩 | | У этого знака не бывает откидной, а сам он лишь трижды проник в середину иероглифа. в остальных 130-ти это правый нижний знак. | 𠂇 𠂇 𠂇 |
| 124 | 卩 | | Похожий на этот знак элемент с откинутой ногой (второй пример) кодируется последовательностью "一 卩". Таковых всего 6 шт. | 𠂇 𠂇 |
| 125 | 卩 | | Похожий на этот знак элемент с оттопыренной ногой (пример 2) кодирует последовательность "丿 卩". | 𠂇 𠂇 |
| 126 | 卩 | | | 𠂇 |
| 127 | 卩 | | | 𠂇 |
| 128 | 𠂇 | | Во всех 20-ти иероглифах с этим знаком он последний. И никакого знакоразличительного смысла от предыдущего нет, одна стилистика написания нижнего правого края. | 𠂇 |
| 129 | 卩 | 卩 卩 卩 | Начались варианты смещения горизонтальных: влево, вправо, в обе стороны. Никакого конструктива не просматривается, хотя где-то, в одном-двух местах, это было важно. Смотрите глазками. Последний пример с добавочной горизонталью, но в right-режиме опознаётся с двух "тыков" и никакая добавочная горизонталь ему не нужна. | 𠂇 𠂇 𠂇 𠂇 𠂇 |
| 130 | 卩 | 卩 卩 | В третьем примере вместо откидной такая же бесполезная скобка. | 𠂇 𠂇 𠂇 𠂇 |
| 131 | 卩 | 卩 卩 卩 | | 𠂇 𠂇 𠂇 𠂇 𠂇 |
| 132 | 卩 | 卩 卩 | Иногда сдвигает горизонтали вправо, а иногда прибавляет справа откидную. | 𠂇 𠂇 𠂇 |
| 133 | 卩 | 卩 卩 | К этому знаку может добавляться одна или две горизонтали как в примерах 2 и 3. | 𠂇 𠂇 𠂇 |
| 134 | 卩 | 卩 | | 𠂇 𠂇 |
| 135 | 卩 | 卩 卩 卩 | "Сборный" знак. Имитирует все "п-образные" элементы с верхним 丿 и дополнительными горизонталями (кроме пустого 卩). | 𠂇 𠂇 𠂇 𠂇 |
| 136 | 𠂇 | 𠂇 𠂇 | Как и предыдущий охватывает все "п-образные" элементы с верхним 丿 и оттопыренной ногой (кроме 𠂇). | 𠂇 𠂇 𠂇 𠂇 |
| 137 | 卩 | | | 𠂇 |
| 138 | 卩 | | | 𠂇 |
| 139 | 𠂇 | | В последней позиции иероглифа, видимо, "для соблюдения принципа", оттопыривает не ногу, а руку. | 𠂇 𠂇 |
| 140 | 𠂇 | | | 𠂇 𠂇 |
| 141 | 𠂇 | | | 𠂇 𠂇 |
| 142 | 𠂇 | | | 𠂇 |
| 143 | 𠂇 | 𠂇 𠂇 𠂇 | Иногда подчёркивается откровенной горизонталью, иногда отмечается еле заметной точкой. Не обращайтесь внимания, в случае необходимости, а это пять-шесть раз на весь набор, вберите вручную. | 𠂇 𠂇 𠂇 𠂇 |
| 144 | 𠂇 | | | 𠂇 |

| № п/п | граф | варианты | описание | примеры |
|-------|------|----------|---|---------|
| 145 | 𠄎 | | Снизу всегда подчёркнут либо короткой горизонталью, либо отмечен точкой, но словари различий в значении не отмечают. Смотрите глазками и выберите нужный. Скорее всего это разнописи. | 纒 纒 |
| 146 | 𠄏 | | | 𠄏 |
| 147 | 𠄐 | | | 𠄐 |
| 148 | 𠄑 | | Осторожно, при реальном вводе основной вариант трудно отличим от предыдущего знака (№147) и от комбинации 𠄑 丨 | 纒 |
| 149 | 𠄒 | | | 𠄒 |
| 150 | 𠄓 | | | 𠄓 |
| 151 | 𠄔 | | | 𠄔 |
| 152 | 𠄕 | | | 𠄕 |
| 153 | 𠄖 | | | 𠄖 |
| 154 | 𠄗 | | У трёх из 50-ти иероглифов вертикаль выше горизонтали, как в первом примере. У остальных горизонталь всегда выше. | 𠄗 𠄘 |
| 155 | 𠄙 | | | 𠄙 |
| 156 | 𠄚 | | | 𠄚 |
| 157 | 𠄛 | | | 𠄛 |
| 158 | 𠄜 | | Осторожно. Легко спутать с комбинацией №50 и №3. См. второй пример. | 𠄜 𠄝 |
| 159 | 𠄞 | | | 𠄞 |
| 160 | 𠄟 | 𠄟 | Не считайте горизонтальные чёрточки. Если их больше двух, всё это знак 𠄟. | 𠄟 𠄠 𠄡 𠄢 |
| 161 | 𠄣 | 𠄣 | Разница вариантов только в высоте третьей вертикали. | 𠄣 𠄤 |
| 162 | 𠄥 | | | 𠄥 |
| 163 | 𠄦 | | | 𠄦 𠄧 |
| 164 | 𠄨 | | | 𠄨 𠄩 |
| 165 | 𠄪 | | | 𠄪 𠄫 |
| 166 | 𠄬 | | | 𠄬 |
| 167 | 𠄭 | | | 𠄭 |
| 168 | 𠄮 | | | 𠄮 |
| 169 | 𠄯 | | | 𠄯 |
| 170 | 𠄰 | | | 𠄰 |
| 171 | 𠄱 | | | 𠄱 |
| 172 | 𠄲 | | | 𠄲 |
| 173 | 𠄳 | 𠄳 | Есть вариант, но разницы никакой нет. Две горизонтали, три горизонтали... В данном случае, это не "диагностический" признак. Можно их и вообще потерять, как в последнем примере. | 𠄳 𠄴 𠄵 |
| 174 | 𠄷 | | | 𠄷 |
| 175 | 𠄸 | | | 𠄸 |
| 176 | 𠄹 | | | 𠄹 |
| 177 | 𠄺 | | | 𠄺 𠄻 |
| 178 | 𠄼 | | | 𠄼 |
| 179 | 𠄽 | | | 𠄽 |

| № п/п | граф | варианты | описание | примеры |
|-------|------|----------|--|-----------|
| 180 | 𠄎 | | | 𠄎 |
| 181 | 𠄏 | 𠄐 𠄑 | Домик всегда имеет крышу, но иногда теряет обе стены (первый пример) или одну (третий пример). | 𠄒 𠄓 𠄔 |
| 182 | 𠄕 | | | 𠄖 |
| 183 | 𠄗 | 𠄘 | В двух случаях выравнивает откидную как в примерах 2 и 3. | 𠄙 𠄚 𠄛 |
| 184 | 𠄜 | | | 𠄝 |
| 185 | 𠄞 | 𠄟 | Из 18 иероглифов с этим знаком в четырёх ² сполз по вертикали вниз (все они в примерах). | 𠄠 𠄡 𠄢 𠄣 𠄤 |
| 186 | 𠄥 | | Этот знак уж очень похож на результат попытки скорописи знака № 183, но встречается и вместе с ним. | 𠄦 |
| 187 | 𠄧 | | | 𠄨 |
| 188 | 𠄩 | | | 𠄪 |
| 189 | 𠄬 | 𠄭 𠄮 | В двух иероглифах знак потерял левый элемент 𠄯 (второй пример). | 𠄰 𠄱 𠄲 |
| 190 | 𠄴 | | Завершающий знак. Никогда не встречается внутри иероглифа. | 𠄵 |
| 191 | 𠄷 | | Также завершающий знак, иногда несёт на себе дополнительную горизонталь. | 𠄸 |
| 192 | 𠄹 | 𠄺 | Ещё один завершающий знак. В двух случаях его две горизонталы сползли вниз. Оба этих варианта в примерах. | 𠄻 𠄼 𠄽 |
| 193 | 𠄿 | | Чаще всего завершающий знак, если после него не следует 𠄿 или 𠄿. Правда, в одном случае он вырвался вперёд (пример 2). Но всегда начинается ушками (𠄿) и завершается крюком вправо (𠄿), поэтому узнаётся всегда. | 𠄿 𠄿 𠄿 𠄿 |
| 194 | 𠄿 | | Только в одном случае из 49-ти после знака 𠄿 стоит ещё нечто (пример 2). | 𠄿 𠄿 |
| 195 | 𠄿 | | | 𠄿 |
| 196 | 𠄿 | | | 𠄿 |
| 197 | 𠄿 | | | 𠄿 |
| 198 | 𠄿 | | Этот обычно завершающий знак иногда вырывается вперёд. | 𠄿 𠄿 |
| 199 | 𠄿 | 𠄿 | Два варианта этого знака: один потерял горизонталь, второй потерял точку. Обратите внимание на отличие от знака 𠄿 - его откидная никогда не пересекает вертикаль. | 𠄿 𠄿 𠄿 |
| 200 | 𠄿 | 𠄿 | В одном иероглифе встретился знак с двумя откидными сверху, в одном случае он вырвался вперёд. Почти в трети случаев (10 из 26) теряет точку как в последнем примере. | 𠄿 𠄿 𠄿 𠄿 |
| 201 | 𠄿 | | Обратите внимание на "тонкие" различия между иероглифами в примере. Скорее всего, просто описка в источнике. Ну, что есть. | 𠄿 𠄿 |
| 202 | 𠄿 | | Этот знак из 22-х случаев только 15 раз оказывается завершающим, но "ногу" всегда протягивает до конца. | 𠄿 𠄿 |
| 203 | 𠄿 | | Этот и следующий знак неожиданно оказались исключительно завершающими, поэтому и попали в последний раздел матрицы. | 𠄿 |
| 204 | 𠄿 | | исключительно завершающий знак | 𠄿 |
| 205 | 𠄿 | | Завершающий знак с одним исключением, когда он каким-то образом попал внутрь. | 𠄿 𠄿 |

