

Деопик Д.В.

### **Соотношение вероятностного и систематического в древнекитайском памятнике (анализ базы данных «Шаньхайцзина»)**

Как используются базы данных, какие возникают проблемы, когда база данных уже сформирована и надо решать какие-то задачи. Дело в том, что создание баз данных, приобретающее все более и более массовый характер, позволяет ставить, а порой и отвечать на вопросы, о которых раньше размышлять не приходилось. С другой стороны сплошь и рядом базы данных создаются некоторым образом вообще, а не под конкретную цель, и это создает определенные трудности владельцам этих баз данных. Был составлен некоторый корпус баз данных по «Шаньхайцзину» (цзюани 1–5, объединенные общим названием «Горы»), древнекитайскому геомантическому памятнику, представляющем собой самый ранний источник о пространственных представлениях древних китайцев, и поставлена была одна конкретная проблема – являются ли представления китайцев V-IV вв. до н.э. о пространстве, отраженные в памятниках этого времени исходными для этой цивилизации, исходными, или им предшествовала некоторая принципиально иная модель предоставления о пространстве. Для этого «Шаньхайцзин–Горы» является единственным источником, так как он считается более ранним. Он бесспорно самый древний, даже при беглом просмотре возникает ощущение, что он принадлежит к несколько иной традиции. Кроме его данных по нескольким векам до середины I тысячелетия до н.э. у нас собственно говоря нет.

Конкретно проблема выглядит следующим образом – были ли представления древних китайцев о пространстве жестко нормированы, как привыкли думать все китайцы и большинство

китаистов, или они были вероятностными, не подчинялись четким схемам, что мы видим в «Ицзине» и круге культурных явлений, с ним связанных. Для ответа на такой вопрос, примеров, конечно, было бы недостаточно. Нужен анализ всего текста. Целью было рассмотреть всю информацию пространственного характера, содержащуюся в «Шанхайцзине-Горы», выделить из нее ту часть, которая системологична, то есть похожа на то, что мы видим в более поздний период и до сегодняшнего дня, выяснить, насколько велика эта часть, как выглядит все остальное. А если выявятся две группы связей – системологические и вероятностные, то надо попытаться установить, в каких сферах по преимуществу мы видим то или другое. Это позволяет представить себе механизм внедрения поздних, как представляется, системологических элементов в ткань достаточно стандартного для общих представлений – раннего человечества о пространстве более древнего памятника вероятностного характера, где все характеристики востока, севера, запада и так далее не являются абсолютными с всегда присущими только им характеристиками, а реализуются с определенными вероятностями. (К примеру, некоторые растения или животные, которые характеризуют по преимуществу запад, могут быть встречены и в других местах). Так выглядит задача.

В связи с этим встал целый ряд проблем, которые возникают при использовании баз данных, о которых мне хотелось кратко рассказать. Но в начале надо немного рассказать о памятнике и связанных с ним текстологических проблемах. Это не рассуждение о пространстве, это список дискретных описаний 450 гор-духов древнего Китая. Напомним, что исследуется первая часть памятника – «Шанхайцзин-Горы». Это памятник, состоящий из списков – наиболее удобный для создания первоначальных баз данных материал. Возникает вопрос, для историка в общем-то достаточно существенный – какова должна быть методика количественного

исследования текста, состоящего в основном из дискретных списков. Тем более, что «Шаньхайцзин-Горы» далеко не единственный, китайская культура далеко не единственная, породившая тексты такого рода. В определенной степени методика исследования списков дискретных описаний является путем исследования любого текста. Автор предлагает положительный ответ на вопрос: сводим ли повествовательный текст любого жанра, кроме художественной литературы, в той или иной форме в системе списков или нет? Этой работе у автора предшествовал анализ по похожей методике первой китайской хроники «Чуньцю». Это хроника и большинство ранних летописей мира, в той мере, в которой автору приходилось их обсчитывать, показывают, что большая часть материалов летописи тоже сводима к спискам. Это – список характеристик войн, список характеристик посольств, список характеристик переворотов. То есть текст в значительной своей части может быть разложен на однотипные высказывания, в которых их основные элементы образуют парадигмы (все виды субъектов, все виды отношений между ними и т.д.), различные значения из которых и формируются списки (список характеристик посольств и т.д.).

Определенные категории текста, важные для исторических реконструкцией, сведены к спискам быть могут. Но степень удаленности естественных текстов, содержащих списки в скрытом виде (летопись, например) от текстов, состоящих из очевидных списков – разная. И поэтому показалось целесообразным заняться памятником «чистой» категории – целиком состоящим из списков и требующих работы по выделению списков (см. исследования автора по «Чуньцю»). Это позволило кое-что выяснить.

Уточняя, что является целью исследования, автор позволит себе несколько банальностей, касающихся работы с базой данных. Общая цель – получение некоторой новой информации, содержащейся в естественном тексте, но в очевидной форме там не представленной. В

данном случае – о системе пространственных представлений, которыми руководствовался автор «Шаньхайцзин-Горы». Необходимым этапом здесь является различие наборов информации о разных частях пространства, т.е. различия разных частей пространства одной от других.

Для этого необходимо прежде всего проанализировать текст на жанровую однородность (по базам данных это делается очень легко). Напомним, что образующие памятник стандартные тексты состоят не только из перечислений тех или иных характерных для точек пространства явлений как материального, так и нематериального порядка, но и из определенного набора объединяющих их в группы итоговых формул. Простейшая работа с базой данных позволяет выявить «сгустки» нетипичных формул и характеристик, которые являются формальным основанием для исключения из основной части исследования этих «сгустков» (для последующего отдельного анализа), если внешний контроль содержания, имеющий отношение к основному тексту, формулировка тоже подтвердит их сущностное сходство между собой. Такими явились здесь концовки описания горных цепей, описаний частей света и самого памятника. Это бесспорно явление иного жанра, оно может быть поздним, может быть синхронным, но бесспорно оно построено на каких-то иных принципах, которые к реконструкции исходного текста могут быть привлечены во вторую очередь. Хотя бы потому, что они содержат некоторые элементы обобщения, а именно эта часть (и особенно в китайской культурной традиции) наиболее подвержена поздним переделкам.

В целом реконструкция древнего видения пространства, в целом и различных его частей была проведена не на отдельных примерах, а на том, как то или иное явление, характеризующее часть пространства реализуется или не реализуется на всем массиве. «На руках» это сделать было бы совершенно невозможно.

В начале несколько примеров того, что дал количественный анализ всего текста. Так, для пространства «Шаньхайцзина-Горы» не присуще деление на север-юг и восток-запад. Тот пресловутый крест, на котором строится якобы исконные представления древних китайцев о пространстве, здесь отсутствует нацело. Противопоставляются друг другу не север – югу, а восток – западу, а противопоставляются север и запад – востоку и югу. Перед нами совершенно необычное виденье мира. Причем, повторяю, это сближение севера с западом, а юга – с востоком постоянно прослеживается на десятках признаков и сотнях их значениях, которые были проанализированы. Прелесть работы с базой данных в том, что мы располагаем всей информацией. Через систематизацию различных признаков описания горы и «цепей» гор и были получены пространственные характеристики сторон света. Была получена надежная информация и о том, что представления о пространстве этого времен принципиально отличаются как от более поздних частей памятника, так и от более поздних представлений древних китайцев. Это – не плоскостное пространство или зональное, как у нас и как у китайцев впоследствии, а линейно-точечное. И исключений из этого нет, практически. Это необычный для нас тип пространства, но он достаточно хорошо прослежен, чтобы вопрос далее не был предметом дискуссий.

Теперь укажем, что нужно было сделать прежде всего в ходе исследования. Нужно было выделить, что достаточно легко, когда все уже сведено в базы данных, разные уровни описания. В тексте представлены в естественном языке, четыре уровня – фразовый, то есть описание горы-духа (горы, являющейся духом) самой по себе («наверху много нефрита» и т.п.), описание цепи гор-духов, описание части света и описание всего комплекса «Шаньхайцзин-Горы» (ранняя часть памятника) в отличие от других частей «Шаньхайцзина». Оценка чистоты классификации исходных блоков

данных памятника, его частей, показал, что наиболее интересными уровни – это цепи гор и сами горы, то есть два наиболее массовых критерия. Вещь достаточно тривиальная, можно было и догадаться, но сама по себе проверка была необходима. Вторично были проверены концовки, которые и поэтому критерию тоже «отвалились» от остального текста. Мало того, когда пошел анализ на этих уровнях, то выделились, тоже достаточно просто и надежно, позже вставленные горные цепи, позже вставленные горы. То есть простейший математический анализ без особых усилий позволил четко противопоставить целый ряд интерполяций или вставок из других памятников (видимо, более поздних) остальному тексту. После этого они выводились из исследования или исследование велось с учетом их нетипичности. Опять же располагая базой данных всегда можно попробовать реализовать несколько вариантов группировки гор-духов, которые раньше приходилось оценивать только приблизительно, отбирать наиболее вероятные. И можно сказать, что исключения эти незначительны, менее 5 % базы; их изъятие делало характеристику всего остального более четкой.

В связи с работой с базой данных в данном случае появилась возможность более серьезно решать проблему, которая до сих пор мешает тем, кто работает с базами данных. Это проблема весов значений признаков; кто с этим работал – все знают, что если веса не определены, то можно дальше уже не трудится. В то же время процедуры решения этой проблемы весьма разнообразны, единых методик не существует. И накапливать опыт в этом отношении необходимо, потому что, если значения признаков признаны условно равноценными, то чем сложнее примененная после этого формула (пример – факторный материал), тем более нелепым будет результат. Пока сейчас в гуманитарных науках работа с базами данных – процесс итерационный, когда результаты одного короткого этапа работы определяют характер последующего шага. И результаты, и этот следующий шаг – понятны историку, текстологу,

сформулированы в языке его науки. Если мы, решая сразу большую часть задачи, не определим веса значений признаков и применим обобщенную сложную формулу – мы не можем сущностно интерпретировать результаты. На какой-то момент мы теряем сущностный контроль над прослеживаемыми нами закономерностями, а потом удивляется результатам и автор, удивляются и читатели, если таковые появятся. А вот тогда, когда появляется возможность быстро «гонять» готовую базу данных в зависимости от результатов предшествующего анализа, начав с решения проблемы таксономии признаков, дело в значительной степени упрощается. Потому, что можно разбить дальнейшую работу на краткие этапы, на каждом из которых мы можем сохранить живую связь с материалом, применить внешний контроль, проверить все варианты корреляций, сделать возможные интерпретации. А затем идти дальше. Хочется еще раз подчеркнуть, что до проведения работы по таксономии признаков все остальное в значительной своей части бессмысленно или остается в пределах интуиции и здравого смысла, когда можно работать и без баз данных.

Как конкретно велся анализ ? На базе основной части текста (без концовок в описании каждой «цепи» гор) строились частные матрицы стандартного облика: слева – «цепи» гор, объединенные в цзюани (книги) по странам света (5 цзюаней: Юг, Запад, Север, Восток, Центр). Сверху: признак и все его значения. В дальнейшем признаки сравнивались по их разделяющей способности, т.е. по таксономической ценности, с одной стороны, и по связи с некоторыми внешними критериями, комментаторским материалом, и устанавливалась классификация «цепей» гор (или цзюаней и их частей) или фиксировалось ее отсутствие и т.д. Это давало определенные результаты. Прежде всего были отсечены редкие признаки, и значения признаков таксономически не существенные, так как не могли характеризовать относительно массовые группы

«гор-духов», отсечены общие для всех объектов значения признаков, которые характеризуют памятник в целом, а не ту или иную его часть. Затем были прослежены корреляции между теми значениями признаков, которые достаточно массовы (три значения одного признака – по 30 % на каждый, условно). Все это удавалось проделать многократно и, фиксируя корреляцию значений одних признаков с значениями других, исследователь одновременно накапливал информацию о специфике частей памятника, которые и являются объектом характеристики. Выяснилось, что памятник состоит не из пяти, а из шести достаточно четких частей, частично не совпадающих с исходным (с оглавлением, грубо говоря) членением памятника. В дальнейшем совпадение или несовпадение с этим членением распределения значений новых привлекающихся признаков служила одним из критериев для определения таксономической ценности тех или иных признаков и их значений (но – не единственным).

В результате прослеженное новое членение подтвердилось. И, что самое главное, были получены надежные, но не стопроцентные, а вероятностные призначные характеристики как для цзюаней, так и для их групп. Одновременно, анализ поведения признаков на массиве памятника позволил определить наиболее легко решающие нашу задачу классификации и выделения интерполяций, признаки и их значения. Не столько минимум их, сколько некоторый оптимум. Это открывает для данного текста возможности, которые были отчасти ранее реализованы на других массивах, прежде всего – массивах документов, на эпиграфических массивах ЮВА, где также велся поиск наиболее существенных для характеристики содержания памятника (таксономически ценных) признаков на группе родственных памятников.

Приведем некоторые соображения о вероятностном или системологическом характере памятника. О самом главном уже сказано, все характеристики – плавающие, вероятностные, а не



стопроцентные. Это значит, что системология в том виде, в котором она дефинирована в наших (и не наших) исследованиях здесь практически не реализуется. А где системология все таки здесь реализовалась ? В одной, но весьма существенной сфере. Это – расстояния между горами. Анализ цифр расстояний между горами позволил сделать ряд наблюдений в привычной для синологов области системологии, хотя эти наблюдения тоже оказались, как показывают беседы с моими коллегами здесь и специалистами их Китая, достаточно нетривиальными. Выяснилось, что расстояния, которые указаны в «Шаньхайцзине-Горы» не являются реальными ни в какой степени и не были ими никогда. Реальность отражена неким причудливым образом. Для центральных «цепей» гор более известных и располагающихся на прародине китайцев – расстояния более короткие и конкретные. Для территорий удаленных – расстояния более крупные и обобщенные. Но и квазиконкретные расстояния в центре кончаются двойкой, пятеркой и их суммой – семеркой. Остальные цифры табуированы, т.е. перед нами не реальность, а определенная магия цифр (кстати, достаточно необычная для привычных для нас реконструкций китайских числовых систем). Есть и табуированные для центра цифры, характеризующие только внешние территории (типа 80). И так далее. Не менее интересно и другое. Две вещи характеризуют «цепи» – это расстояния между горами и ориентация. И вот мы видим в центре, что чем дальше от реального центра проживания древних китайцев, где все формировалось, тем расстояния более крупные, более обобщенные и тем меньше конкретных извилин, которые характеризуют «цепь» (т.е. имевших сходство с реальной горной цепью, в отличие от унылых «горных» линий периферии, уходящих по прямой линии неведомо куда). Это говорит о том, что простая и естественная ситуация, возникающая когда люди думают о пространстве (больше четкость представлений о близком к себе и

более расплывчатое – о далеком) имеется и здесь, но в преобразованной, условной форме. Причем это, повторяю, не сфера дискуссии, поскольку это легко прочитывается, легко демонстрируется. А говорит это очень о многом. Здесь, как и при характеристике частей пространства, где очень редко характерное значение признака реализуется в данной части всеми ста процентами своих реализаций, действуют вероятностные законы, действуют общие расплывчатые законы представления о пространстве, но никак не жесткие законы системологии, которые в это время, видимо, не были созданы и которыми не руководствовались применительно к представлениям о пространстве. О другом нам судить пока трудно.

Деопик. Д.В. Проблемы методики исследования эпиграфического комплекса применительно к задачам социально-экономического анализа (на материале бирманской эпиграфики), - Вестник Московского университета, Востоковедение. М., 1977, № 2

Деопик. Д.В. Опыт количественного анализа древней восточной летописи "Чуньцю". - Математические методы в историко-экономических и историко-культурных исследованиях. М., 1977.

Деопик. Д.В. О системе пространственных представлений древних китайцев (По данным "Шань хай цзин"), 9-ая Научная конференция "Общество и государство в Китае". Тезисы докладов. Ч. 2. М., 1978

Деопик, Гегемония и гегемоны по данным "Чуньцю", - Общество и государство в Китае, М., 1978

Деопик Д.В., Кулланда С.В. Простейшие признаки яванского эпиграфического массива VII - начала X в. как источник по истории раннесредневековой Явы, - Этническая история народов Восточной и Юго-Восточной Азии в древности и средние века. М., 1981

Деопик. Д.В. Вопросы количественного анализа эпиграфических памятников, - Вестник Московского университета, Востоковедение. М., 1983, № 4

Деопик Д.В. Исторический контент-анализ, - Комплексные методы в исторических исследованиях: тезисы докладов и сообщений научного совещания. М., 1987

Опубликовано в Базы данных по истории Евразии в средние века. Выпуск 4-5. М. 1997, с. 57-65