SPATIO-TEMPORAL DATA MODEL FOR INTEGRATING EVOLVING NATION-LEVEL DATASETS

Alexandre Sorokine[a,*], Robert N Stewart[a]

[a]Oak Ridge National Laboratory, (SorokinA,StewartRN)@ornl.gov

KEY WORDS: data model, big data, spatiotemporal databases

ABSTRACT:

Ability to easily combine the data from diverse sources in a single analytical workflow is one of the greatest promises of the Big Data technologies. However, such integration is often challenging as datasets originate from different vendors, governments, and research communities that results in multiple incompatibilities including data representations, formats, and semantics. Semantics differences are hardest to handle: different communities often use different attribute definitions and associate the records with different sets of evolving geographic entities. Analysis of global socioeconomic variables across multiple datasets over prolonged time is often complicated by the difference in how boundaries and histories of countries or other geographic entities are represented. Here we propose an event-based data model for depicting and tracking histories of evolving geographic units (countries, provinces, etc.) and their representations in disparate data. The model addresses the semantic challenge of preserving identity of geographic entities over time by defining criteria for the entity existence, a set of events that may affect its existence, and rules for mapping between different representations (datasets). Proposed model is used for maintaining an evolving compound database of global socioeconomic and environmental data harvested from multiple sources. Practical implementation of our model is demonstrated using PostgreSQL object-relational database with the use of temporal, geospatial, and NoSQL database extensions.

1. INTRODUCTION

Most global economic and environmental data supplied by governments and international organizations are reported at the national level. At the first glance such datasets may seem to be trivially conceptualized as a rectangular three-dimensional matrices where columns represent variable names, rows represent countries, and time-stamped variable values are represented as slices in the third dimension. Such a data structure, if it existed, would be well suited for data mining methods like, for example, factor analysis or Candecomp decomposition or other techniques. However, real-world nation-level data always need additional processing before it can be converted into a form suitable for data mining algorithms.

The first difficulty is that countries appear, disappear, merge or split. For the three decades examples of the countries that once existed and then disappeared or newly appeared include USSR, Yugoslavia, and Eritrea just to name a few. In a three-dimensional matrix representation emergence or disappearance of countries introduces a number of problems. First, emerging or disappearing countries result in a large number of missing values. These missing values have to be imputed before the processing is done because only a few data mining algorithms are able to handle large amounts of missing data. Second, missing values represent a special kind of data abnormality. Typically, missing values result from the observations are not being collected or reported. Such cases are easily handled by various imputation methods given enough data points at hand. In the case when a country emerges or disappears, standard imputation methods may not produce reliable results and may have problems in interpretation.

The problem becomes more challenging when the data have to be merged from multiple sources over prolonged periods of time like 25 years or more. During such time intervals, the number of countries experiencing transformation events (*e.g.*, new countries appear, existing countries disappear, countries gain or lose territory) is big enough to have significant impact on global indicators. For example, since 1989 the Central Intelligence Agency World Factbook (WFB, Central Intelligence Agency, 2013) contains about 80 cases when a country was affected by changes in its borders, its international status, or has appeared or disappeared from the world map. The cumulative effect of such events on global indicators is shown in Table 1. In addition, large number of such transformations occur in the unstable regions that are of high interest to the researchers.

Indicator	Value	% of global
Number of affected countries	62	25%
without minor areas	40	16%
land area	24 mil sq km	17%
population	750 mil	10%

Table 1. Effects of transformation events on global indicators in 1989–2011 (calculated based on the CIA World Factbook, 2013)

The other difficulty is that global databases created by different governments, academic, or international organizations view the sets of currently or previously existing countries and their histories differently and store their data according to their views. One reason for such discrepancies is that not all countries are recognized by the government of all other countries. For example, Serbia does not recognize Kosovo or Israel does not recognize the Palestinian State (as of 2017). The same problem arises when international treaties are

interpreted differently by different governments or organizations. In addition, the datasets are frequently updated and corrected. Updated versions of the datasets do not only add new data but often contain corrections to the sets of the countries and sometimes retroactively change country histories.

Data mining is not the only area that requires tracking of country identities and histories. For example, a number of projects are investigating how involvement in international treaties and alliances affect country well-being and world piece (Singer, 1988, Li et al., 2017, others). Such analysis requires a very precise tracking of country histories and their involvement in alliances. There is a plethora of other applications that will benefit from good understanding of country histories and their perceptions by different communities.

This paper presents some results of a study to develop a well-founded approach to handling national- and subnational-level data based on the identity and country histories. Proposed framework has been implemented in a practical datamining system called W-STAMP (World Spatiotemporal Analytics and Mapping Project). Currently W-STAMP contains 15,000+ attributes sourced from more than a dozen public global datasets (Stewart et al., 2015). These data consist of close to 18 million records that characterize more than 250 world entities over approximately 50 year period. The proposed framework is able to accommodate real-world dynamics and changes in the datasets with the purpose of integrating multiple sources of data in a single analytical workflow.

2. FOUNDATIONS

In the datasets that fall within the scope of this study all data records are linked to entities at the national or subnational levels. In different data products, such entities are called differently. For example, in the WFB they are called "world entities", in the World Bank and UNEP data they are referred as countries, and in the Correlates of War dataset (Correlates of War Project, 2011) in most cases they are called "states". It is quite clear that different data vendors may have different understandings of such entities both at conceptual level and at the level of specific entities listed in the datasets.

Literature analysis shows that there is no simple or unambiguous definition of a "state" or a "country" and often the names of the states or countries (e.g., USA, Germany, England) are used in multiple senses. The meaning of the term "state" has been studied in (Robinson, 2012, Robinson, 2010, Crawford, 2006). There is a vast literature in political geography offering a large number of definitions and notions the "state". A "state" can mean a geographical location, a legal entity, a government, a set of institutions, etc. The core to the comprehension of the ontological status of the state is its understanding as a part of social reality and result of social construction (Searle, 1995). This fits well with the challenges of this study to develop an architecture that would be able to handle variations in the understanding of "states" from multiple independent agents. The datasets used in this study are created by different government, international, academic, commercial, non-governmental organizations or activists each having its own concept of state that is adapted to the goals and purposes of those people or organizations. In this study we will use the WFB term *world entity* in place of such words as "state", "country", "territory", and other similar

terms encountered in global datasets. We will use the term *perspective* to designate a set of world entities and their histories that can be deduced from a specific dataset or dataset version. The intention of a *perspective* is to capture the view of a specific data vendor or other agent.

The next question within our framework is the identity of a world entity and preservation of the identity through transformations such as the ones that occur during country breakups, mergers, and other events. A proper handling of identity is important in the context of data mining because it is necessary for tracking changes that happen to single entity. For example, one of the typical questions that arise when a world entity splits into several entities, is whether one of the successor entity preserves the identity of the original one. There is no single answer to this question and different cognizing agents (government or international organization, dataset vendors or media) can understand preservation of a country identity differently and will collect and report their data according to their understanding.

Understanding a world entity, a country or a state as an element of social reality and recognizing existence of multiple cognizing agents, we have set our goal to build a system that will be able to support multiple representations of world entities and their histories and that will have the means to reconcile the entities and their histories between different representations. Our generalized model for world entities is based on the following assumptions: (1) each world entity has an identity, (2) the life of each world entity has a beginning and has an end but exact time of them are not necessarily known, (3) a world entity is not reducible to its territory, *i.e.*, it may preserve identity even it loses or gains territory.

The concept of identity of geographic objects and its change has been investigated in numerous publications and standards (Stehle and Peuquet, 2015, Claramunt and Theriault, 1996, Doerr, 2003, Worboys, 2005, others). In case of the proposed framework we need to represent world entities and events that affect their existence. For such purpose event-oriented architecture is a natural choice. Event-oriented architecture stores changes explicitly by recording events that have to be interpreted relatively to the initial state of the data (Yuan, 1999, Worboys, 2005).

Another common approach is representing the data as time stamped slices. This approach is used in most of the datasets where we have sourced the data from. A number of timestamping approaches have been described in literature. Timestamps can be attached to various database objects like tables, tuples, or specific cell values (Yuan and McIntosh, 2002). All of these approaches have been used for temporal geospatial information and each of them has its own advantages or disadvantages. For example, timestamping the whole table or database is convenient when a user is interested in specific time slices but results in a large number of duplicated or missing values when the database is updated from multiple sources.

Event-based approach can capture change precisely and is used by some information systems that store the data on the history of administrative units (Gantner et al., 2013, Gantner, 2011, Goodwin et al., 2008, Plumejeaud et al., 2009). The disadvantage of the event-based approach is that in such systems certain type of queries are more difficult to formulate and they require more computational resources.

Practical examples of similar spatio-temporal databases can be found in a number of studies related to the history of administrative units within the countries (Gantner et al., 2013, Gantner, 2011, Plumejeaud et al., 2009, Frank et al., 2003, Goodwin et al., 2008). The history of administrative units may seem similar to the history of world entities but there is at least one important difference. In the case of a single country, the boundaries and histories of its administrative units are defined by its government. However, in the case of global geopolitical data, there is no single authority that defines existence and records histories of the world entities. According to our knowledge, there is still a gap in understanding how spatio-temporal information from such disparate sources can be integrated. Our proposed solution is outlined in the remainder of the paper.

3. MATERIALS

The global data that were used in this study have been sourced from multiple publicly available datasets. Two of these datasets (The World Factbook and Correlates of War) were used to compile historical event information. The World Factbook (WFB) is a reference dataset developed by the Central Intelligence Agency for US policymakers and general public (Central Intelligence Agency, 2013). WFB contains information of around 250 world entities mostly pertaining nation states that are characterized using about 250 variables. One of the advantages of WFB is that its information is in public domain and its availability has prompted a large number of research projects and derivative data products. Digital form of WFB is available starting from 1989 until present.

In terms of spatio-temporal data model, the WFB relies on a simple time-slice approach with a standard set of variables provided for each year per each country. Since 1989 there were several additions to the set of variables like, for example, inclusion of variables depicting Internet penetration and electronic communications in 2001. WFB is updated annually with more frequent updates planned in the near future.

The Correlates of War (COW) dataset has been in development by the academic community since 1960-s with the goal of studying the dynamics of war and international conflict and determining factors that can cause wars and conflicts (Singer, 1988, Correlates of War Project, 2011). COW contains data at the state level pertaining to the existence of states, conflicts, wars, international treaties, and a number of variables characterizing economic development, military, conflict, trade, and others (Tir et al., 1998, Sarkees and Schafer, 2000). In terms of the spatio-temporal data model, COW combines time slices and events. In this dataset, each country is assigned a unique code that links together several tables containing state lifelines, state system membership, and other data like religious adherence, disputes, material capabilities, etc. In addition, there is a downloadable[1] set of ESRI shapefiles that contain time-indexed sea and land boundaries of major countries (Weidmann et al., 2010). COW is updated at irregular intervals every few years.

The data used in this study represent a typical example of the global information at the nation-state level produced by different data vendors. These data describe the same domain but have different perspectives on it and there is a number of differences in how coun-

tries are represented. For example, WFB contains the data on Yugoslavia from 1989 until 1991 and then indicates its breakup. In COW dataset Yugoslavia is present until 2011.

Additional data have been sourced from the World Bank Open Data[2], Global Health Observatory of the World Health Organization[3] and a several others. These datasets are created and maintained by large international organizations and contain predominantly attribute information. None of them track country histories but instead interpolate existing world entities back in time. One of the newly emerged projects that was used in this study is Thenmap[4]. Thenmap is a community repository of historical national and subnational boundaries. The data are available via a web API and represent borders between the territorial units with additional metadata.

4. METHOD: CONCEPTUAL FRAMEWORK

The proposed framework is defined as a set of event that determine existence of the world entities and bound their lifelines, record territory exchanges between them, and link attribute values to the countries and time stamps. We define several event types that are depicted in Table 2. Event types are distinguished by their function, the number of world entities that precede and succeed the event (predecessors and successors), and the fate of the predecessors and successors. The first group of the events are existential events that occur when a country emerges or ceases to exist. This may happen as a result of secession from a predecessor country, breakup of a country, or merger of several countries. There are four main event types in this group numbered 1 to 4 (E_1, E_2, E_3, and E_4):

- In case of a E_1 event type predecessor world entity ceases to exist and several new world entities emerge. The territory of the predecessors is distributed among its successors. Examples of such event are breakups of the USSR and SFRY in the early 1990s.

- In the event of E_2 several predecessor world entities cease to exist and a single new world entity appears. Unification of Germany can be represented as an E_2 event.

- E_3 is similar to E_1 but the predecessor continues to exist after experiencing loss of its part to the new world entity. Separation of Eritrea from Ethiopia is recorded as such event.

- E_4 requires two or more world entities with one of the entities continuing its existence and incorporating territories of all of its predecessors.

There are always multiple world entities that participate in a single existential event. As a result of an event a world entity may emerge, cease to exist or continue existence. The fate of an entity in an event is recorded as a special attribute in the database as "start", "end", or "continue".

Changes of a world entity territory are recorded using T_1 events. T_1 event has the same number of successors as the number of predecessors and does not cause any countries to emerge or cease to

---

[1] http://nils.weidmann.ws/projects/cshapes

[2] http://data.worldbank.org/
[3] http://www.who.int/gho/database/en/
[4] http://thenmap.net

Event Type	Predecessors			Successors			Diagram	Example/Details
	min	max	life	min	max	life		
E_1	1	1	end	2	∞	start		breakup of USSR, SFRY
E_2	2	∞	end	1	1	start		unification of Germany
E_3	1	1	continue	1	∞	start		independence of Kosovo from Serbia
E_4	1	∞	end	1	1	continue		French colonies become departments
$E_?$	–	–	end	–	–	start		details unknown
P_1	0	0	–	1	1	start		the earliest record of an entity
P_2	1	1	end	0	0	–		the latest record of an entity
T_1	1	∞	continue	2	∞	continue		territory exchange between RF and PRC
O_1	1	1	continue	1	1	continue		observation

on the diagrams a cross arrow tip (⤪) indicates the end of life of the predecessor and a solid dot (●→) indicates formation of a new entity after the event, the line crossing the event indicates that an entity survives the event, lines with question marks (?) indicate that entity existence is unknown

Table 2. Event Types

exist. A T_1 event may have a corresponding record with the updated country boundaries in the territories table if such information is available. Unlike some other data sources (*e.g., COW*) we do not record if there was a conflict when certain event occurred and if such information is needed it should be associated with a specific event instance as an attribute.

To handle uncertainty inherent in global spatio-temporal data we introduce a pair of events that indicate when we have the knowledge about existence of a particular world entity. P_1 is a special type of an event that is created when a country is mentioned in the database for the first time. It is similar to existential events as it bounds world entity existence in the dataset. P_2 is a converse of P_1 and designates the last time when a specific world entity was mentioned in the dataset. However, neither of these two events designate a start or an end of a world entity existence in the real world but rather they indicate the lack of knowledge about the country existence prior to or after the timestamp of the event. Typically, P_1 event coincides with the earliest data record involving a world entity and P_2 coincides with the time when the dataset had been published. In the case of P_1 and P_2 events the entity fate is interpreted as a database-level event and indicates that information about entity existence is not complete. Queries about existence of an entity before P_1 or after P_2 report its existence as unknown.

Also for some datasets that lack explicitly stored information about country histories we use a special kind of an existential event called $E_?$. For this event type, we only record the fate of the entity without such details as other entities participating in the same event. $E_?$ indicates that information is incomplete and instructs the software not to include the entity into further processing.

We use the same event system to store the values of attributes pertaining to specific world entities. Attribute information is linked into the rest of the framework by the way of an observation events (O_1) using the same event-based model. Such events do not imply any changes to the world entity existence but reuse the mechanism of the existential events. It is important to notice that Table 2 represents only event types that we have encountered in the data source used in this study. Other event types can be added in the future if a need emerges.

One of the biggest challenges in our approach is building the list of transformation events. Only few data sources provide this information explicitly. In this study, such information was available in COW in the form that needed to be adapted to our representation. In case of other data sources it has to be extracted manually from verbal descriptions and implied assumptions. One of the assumptions that we have used is that the presence of a data record pertaining the entity typically indicates its existence. After such information is extracted it has to be verified to be consistent within itself because there may be errors both in the original data and resulting from the extraction process. Automated procedures were developed to ensure consistency of the country histories with the model presented here and avoiding a multitude of common problems like presence of data records when a world entity is not in existence.

To implement the checks we have developed a model of state transitions shown on Fig. 1a. In our system, the world entity can be in one of the three states: E — exists, N — does not exist, and U — existence unknown. The entity can transition from state U to the

state E via a P_1 event or back via P_2 event. Transitions from the state E to state N happen via events of types E_1, E_2, or E_4. The reverse transition from N to E is restricted to events E_1, E_2, and E_3. Participation of an entity in the events of types E_3, and E_4 keeps it in the E state. Events of non-existential types T_1, O_1 do not affect world entity existence and can only occur in E state.
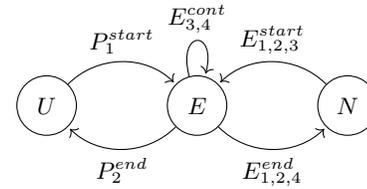
(a) Proposed model
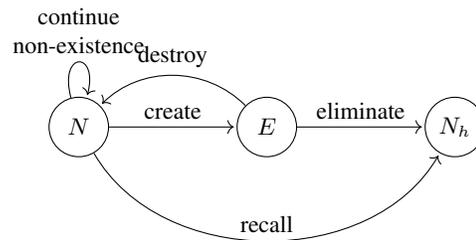
(b) The model proposed in (Hornsby and Egenhofer, 2000, Fig. 8, pg 218)

Figure 1. State diagram

The checks are performed by comparing event types in the pairs of subsequent events against a table of permitted and prohibited transitions (Fig. 2). The first column in Fig. 2 shows the initial state of an entity that transitions to the subsequent state (columns U, E, and N). The cells of the table show the event types such transition may happen through. Symbol $\emptyset$ indicates that such transition is not permitted.

Other checks include: (1) verification that there no events of type O and T when the country is in the states N or U, (2) there are no events of any kind that precede P_1, (3) there are no events of any kind after P_2, (4) the number of successors, predecessors, and survivors is restricted to the values permitted in Table 2.

current state ↗	subsequent state		
	U	E	N
U	$\emptyset$	P_1^{start}	$\emptyset$
E	P_2^{end}	$T_1, O_1, E_{3,4}^{cont}$	$E_{1,2,4}^{end}$
N	$\emptyset$	$E_{1,2,3}^{start}$	$\emptyset$

Figure 2. Prohibited and permitted transitions

The proposed approach has lots of similarities with the model described in (Hornsby and Egenhofer, 2000). In both models entities preserve their identity through their lives. The entities can emerge or cease to exist resulting from events or "operations" in the terminology of (Hornsby and Egenhofer, 2000). In (Hornsby and Egenhofer, 2000) an object (entity) can be found in three states: object does not exist (N), object exists (E), and object is non-existing with history (N_h). Possible transitions are shown on Fig. 1b. However, the model proposed in (Hornsby and Egenhofer, 2000) is not

completely applicable to our domain. First, we need facilities to represent unknown states (U on Fig. 1a). The U state is very important in our domain because most of the countries exist much longer than the time periods for which the data are collected. Also, many databases are not regularly updated thus leaving the question of entity existence between the publication and current date open. Second, we did not find in our domain any cases for "Exists with history" state and transitions to/from it. Our model provides the following advantages over other approaches: (1) it defines a set of states specifically crafted for global socioeconomic datasets, (2) it supports uncertainty inherent in such datasets by introducing the U state, (3) it has a mechanism to handle multiple perspectives at once.

Using the mechanisms of events, we were able to reconcile perspectives between different data sources. Reconciliation works by matching world entities in one perspectives to the entities in another perspective and specifying time intervals when such equivalence is valid for each entity.

5. RESULTS: IMPLEMENTATION EXAMPLE

In W-STAMP our framework is used as a foundation for the analytical system database to track updates in the incoming information, clean the data of erroneous records, and reconcile data sources (Stewart et al., 2015). The users can view country histories and event lists for select regions and time periods. This helps to detect analytical artifacts resulting from changes in attribute values due to changing country territories rather than underlying processes.

W-STAMP utilizes an object-relational database system to store its data with the support for our spatiotemporal framework implemented on top of it. The choice of an object-relational database was guided by multiple considerations. First, most of the existing analytical software have good interfaces with relational database systems. Second, there are good theoretical foundations for handling temporal data types and representations in the database domain. Major concepts like validity time were developed in early 1990-s (Ozsoyoglu and Snodgrass, 1995, Jensen et al., 1994) and fit well with the kinds of objects represented in our study. In a pure relational model, time can be supported by a special temporal data type but special temporal operations are not available. Object-relational model extends time data type with its own sets of operations that can be effectively employed in our system (Stonebraker and Moore, 1996). It is important to note that proposed framework is not dependent upon object-relational database and can be implemented on top of triple stores or graph databases.

In this project we are using PostgreSQL. Most of the temporal logic that is necessary for our framework (predicates like *before*, *after*, *overlaps*, etc.) can be implemented using PostgreSQL data type *time range* that is available in the PostgreSQL core system. For spatial data and operations, we use the functionality of the Post-GIS extension. It is important to notice that our framework can be implemented using other underlying technologies like triple stores or NoSQL if they have necessary facilities for temporal and spatial operations.

Our database representation is designed as a combination of the three-domain model with event-based model described in (Worboys, 2005, Yuan, 1999). Simplified database schema is shown

on Fig 3. We will use `teletype` font to designate the names of database objects.

The table `perspectives` contains a list of perspectives identified by mnemonic codes (*i.e.*, WFB or COW) and their descriptions. The table `world_entities` provides a catalog of all world entities encountered in all perspectives. Each world entitiy is distinguished by its identifier (typically a country name) and is linked to a perspective. `world_entities` table does not contain any timestamps, geometries, or observations. All temporal information is contained in the `timestamp` column of the `events` table. Each event has a type and also is associated with one of the perspectives that can be looked up in the `perspectives` table. Table `transformations` details existential events and holds information about the event types and world entities that are affected by such events. The table `territories` contains geometric information that is linked to the `transformations` table via identifiers of an event and of a world entity. The values of the variables are stored in the `observations` table that is linked via `event_id` to the `timestamp` column in the `events` table. All table are connected by referential integrity constraints that prevent storing of the inconsistent information in the database.

To relate countries in different perspectives we use a table of links between the perspectives (Table 3). The table establishes correspondence between country identifiers (columns `world_entity_1` and `world_entity_2`) in different perspectives (columns `perspective_1` and `perspective_2`) and time intervals when such equivalence is valid (column `validity_interval`). The table is used by the stored procedures that perform such operations as comparison of country sets and country histories between perspectives and fusing variable values from different perspectives. This functionality can be also used to investigate compatibility among perspectives and consistency of country histories within and across perspectives.

6. CONCLUSIONS AND FUTURE WORK

Integrating geospatial and temporal information is always challenging especially if the information comes from disparate sources. In our study, we have investigated multiple global datasets with the goal of merging their data for analysis with the help of data mining and pattern detection algorithms. As a result of this study, we were able to elucidate the structure of spatio-temporal representations used by these datasets and develop a novel database architecture. Proposed architecture is intended to handle multiple spatio-temporal representations (perspectives) of the entities pertaining to the same knowledge domain. The architecture was implemented on top of an object-relational database in a practical data mining application.

Currently we are working on the extending of the proposed framework to incorporate other global datasets that may have their own perspectives. We also plan to add support for other types of information like data on international relations and country participation in treaties and international organizations. We hope that in addition to data mining, the proposed framework can be applied in other areas such as disambiguation of geographic names and for similar tasks related to processing of natural language and social network feeds. In the future we plan to release a working open-source implementation of the proposed approach and extend it to support Linked Data and connect to other systems like DBpedia.
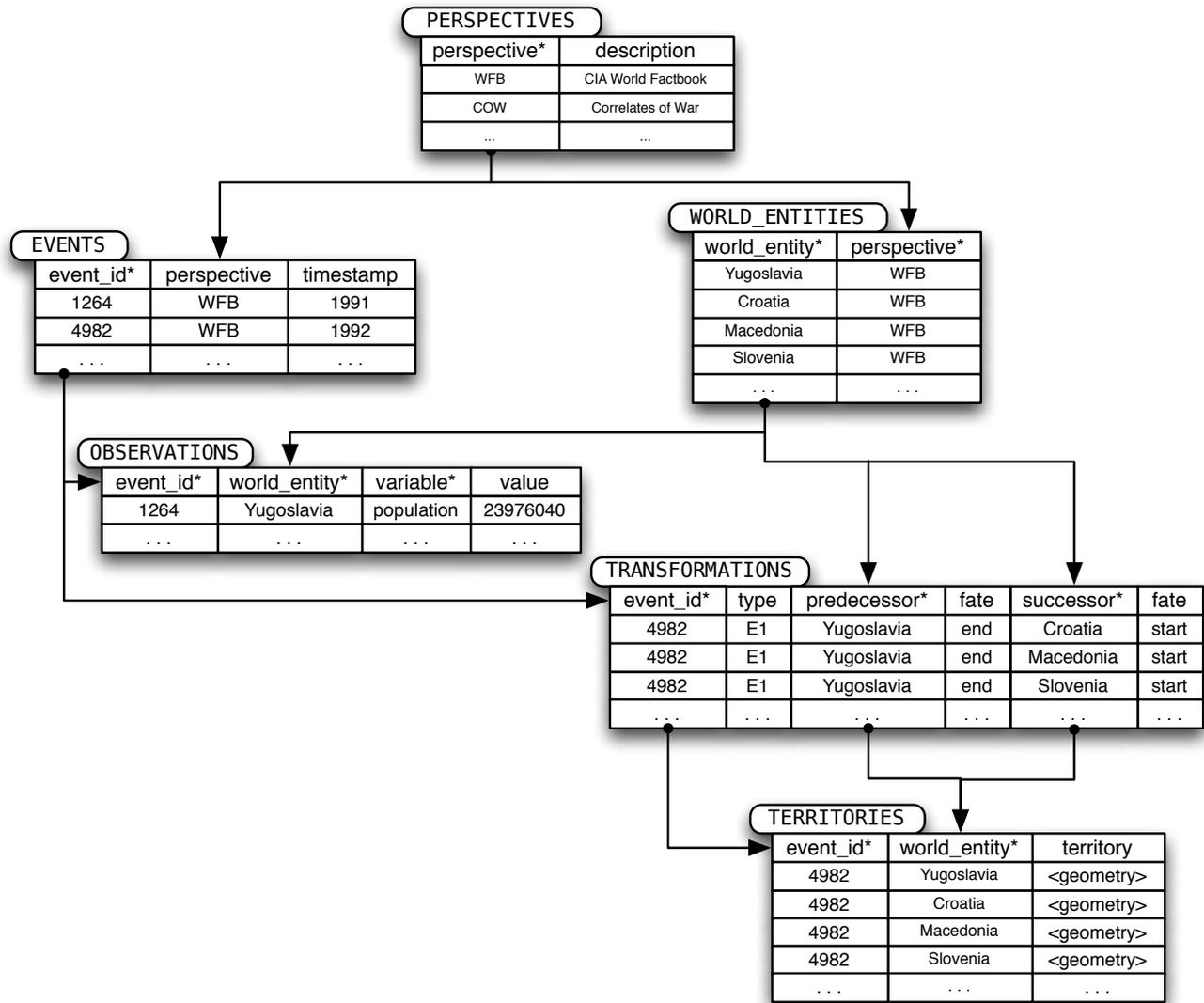
asterisk (*) after the column name indicates a unique primary key

Figure 3. Database Schema (simplified)

perspective_1	world_entity_1	validity_interval	perspective_2	world_entity_2
WFB	Soviet Union	1989–1991	COW	Russia
WFB	Russia	1992–2014	COW	Russia
…	…	…	…	…

Table 3. Relations between *world_entities* in different perspectives

ACKNOWLEDGEMENTS

REFERENCES

Central Intelligence Agency, 2013. *The World Factbook 2013-14*. Washington, DC.

Claramunt, C. and Theriault, M., 1996. Toward semantics for modelling spatio-temporal processes within GIS. *Advances in GIS Research I* pp. 27–43.

Correlates of War Project, 2011. *State System Membership List*. online. `http://correlatesofwar.org`.

Crawford, J., 2006. *The creation of states in international law*. Clarendon Press ; Oxford University Press, Oxford; New York.

Doerr, M., 2003. The CIDOC conceptual reference module: an ontological approach to semantic interoperability of metadata. *AI magazine* 24(3), pp. 75.

Frank, A., Raper, J. and Cheylan, J. P., 2003. *Life and Motion of Socio-Economic Units: GISDATA*. CRC Press.

Gantner, F., 2011. A Spatiotemporal Ontology for the Administrative Units of Switzerland. PhD thesis, University of Zurich, Department of Geography, Zurich, Switzerland.

Gantner, F., Waldvogel, B., Meile, R. and Laube, P., 2013. The basic formal ontology as a reference framework for modeling the evolution of administrative units. *Transactions in GIS* 17(2), pp. 206–226.

Goodwin, J., Dolbear, C. and Hart, G., 2008. Geographical linked data: The administrative geography of Great Britain on the semantic web. *Transactions in GIS* 12(s1), pp. 19–30.

Hornsby, K. and Egenhofer, M. J., 2000. Identity-based change: a foundation for spatio-temporal knowledge representation. *International Journal of Geographical Information Science* 14(3), pp. 207–224.

Jensen, C. S., Soo, M. D. and Snodgrass, R. T., 1994. Unifying temporal data models via a conceptual model. *Information Systems* 19(7), pp. 513—547.

Li, W., Bradshaw, A. E., Clary, C. B. and Cranmer, S. J., 2017. A three-degree horizon of peace in the military alliance network. *Science Advances* 3(3), pp. e1601895.

Ozsoyoglu, G. and Snodgrass, R., 1995. Temporal and real-time databases: a survey. *IEEE Transactions on Knowledge and Data Engineering* 7(4), pp. 513—532.

Plumejeaud, C., Moisuc, B., Bimonte, S., Villanova, M. and Gensel, J., 2009. An object-oriented model for the sustainable management of evolving spatio-temporal information. In: B. Murgante, G. Borruso and A. Lapucci (eds), *Geocomputation and Urban Planning*, Studies in Computational Intelligence, Springer Berlin Heidelberg, pp. 167—186.

Robinson, E. H., 2010. An ontological analysis of states: Organizations vs. legal persons. *Applied Ontology* 5(2), pp. 109–125.

Robinson, E. H., 2012. Reexamining fiat, bona fide and force dynamic boundaries for geopolitical entities and their placement in DOLCE. *Applied Ontology* 7(1), pp. 93—108.

Sarkees, M. R. and Schafer, P., 2000. The correlates of war data on war: An update to 1997. *Conflict Management and Peace Science* 18(1), pp. 123–144.

Searle, J. R., 1995. *The construction of social reality*. Free Press, New York.

Singer, J. D., 1988. Reconstructing the correlates of war dataset on material capabilities of states, 1816-1985. *International Interactions* 14(2), pp. 115–132.

Stehle, S. and Peuquet, D. J., 2015. Analyzing Spatio-Temporal Patterns and Their Evolution via Sequence Alignment. *Spatial Cognition & Computation* 15(2), pp. 68–85.

Stewart, R., Piburn, J., Sorokine, A., Myers, A., Moehl, J. and White, D., 2015. World spatiotemporal analytics and mapping project (WSTAMP): Discovering, exploring, and mapping spatiotemporal patterns across the World's largest open soruce data sets. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 2(4), pp. 95.

Stonebraker, M. and Moore, D., 1996. *Object-Relational DBMSS: The Next Great Wave*. Morgan Kaufmann.

Tir, J., Schafer, P., Diehl, P. F. and Goertz, G., 1998. Territorial changes, 1816-1996: Procedures and data. *Conflict Management and Peace Science* 16(1), pp. 89–97.

Weidmann, N. B., Kuse, D. and Gleditsch, K. S., 2010. The geography of the international system: The CShapes dataset. *International Interactions* 36(1), pp. 86—106.

Worboys, M., 2005. Event-oriented approaches to geographic phenomena. *International Journal of Geographical Information Science* 19(1), pp. 1—28.

Yuan, M., 1999. Use of a three-domain representation to enhance GIS support for complex spatiotemporal queries. *Transactions in GIS* 3(2), pp. 137–159.

Yuan, M. and McIntosh, J., 2002. A typology of spatiotemporal information queries. *Kluwer International Series in Engineering and Computer Science* pp. 63–82.