

## GeoSQL: um ambiente online para aprendizado de SQL com extensões espaciais

Anderson L. S. Freitas<sup>1</sup>, Clodoveu A. Davis Jr.<sup>1</sup>, Thompson M. Filgueiras<sup>1</sup>

<sup>1</sup> Departamento de Ciência da Computação – Universidade Federal de Minas Gerais (UFMG)  
Caixa Postal 702 – 30.123-970 – Belo Horizonte – MG – Brasil

{alsr, clodoveu, thom}@dcc.ufmg.br

**Abstract.** *The standardization of the Structured Query Language (SQL) was important for the popularity of relational databases. The SQL learning process is supported by several resources, such as simulators and interactive tools. However, as far as we could ascertain, there are no tools to promote the learning of spatial extensions to SQL in geographic databases. This work presents GeoSQL, an online learning environment for SQL with spatial extensions, which is able to render the query results and can overlay the results of multiple queries for visual comparison. GeoSQL was developed entirely using free software, and can be operated using only a standard browser.*

**Resumo.** *A padronização da Structured Query Language (SQL) foi importante para a popularização dos bancos de dados relacionais. O aprendizado de SQL é apoiado por diversos recursos tecnológicos, como simuladores e ferramentas interativas. No entanto, não encontramos tais recursos para promover o ensino de extensões espaciais de SQL em bancos de dados geográficos. Este trabalho apresenta o GeoSQL, ambiente online de aprendizado de SQL com extensões espaciais, capaz de renderizar e de sobrepor os resultados de múltiplas consultas para comparação visual. O GeoSQL foi totalmente desenvolvido usando software livre e pode ser operado usando apenas um navegador padrão.*

### 1. Introdução

Gerenciadores de bancos de dados relacionais se tornaram amplamente populares por diversos motivos, dentre os quais se destaca a adoção da *Structured Query Language* (SQL) como linguagem padrão de consulta. Como consequência, o ensino de SQL é parte importante do conteúdo de cursos e disciplinas de bancos de dados em todo o mundo.

O mercado editorial dispõe de uma grande quantidade e variedade de materiais para o ensino de bancos de dados, muitos deles voltados especificamente para a linguagem SQL. Isso reflete a ampla utilização da linguagem no mercado, mas também indica uma potencial dificuldade para o aprendizado. O fato de SQL ser uma linguagem declarativa, e não procedural, requer que o estudante aprenda a pensar dentro da lógica de conjuntos, em vez de algoritmos (Sadiq and Orłowska, 2004). Por isso, atividades práticas individuais são muito importantes.

O uso de SQL como linguagem para acesso a dados geográficos já foi alvo de críticas e restrições (Egenhofer, 1992), porém a definição dos padrões do *Open Geospatial Consortium* (OGC) para representação geográfica em ambientes relacionais (Percivall, 2003), e a evolução dos sistemas de gerenciamento de bancos de dados (SGBD)

objeto-relacionais terminaram por estabelecer um paradigma vitorioso, hoje implementado (embora com variações de sintaxe) em diversos SGBDs, tais como Oracle e outros de código livre. Apesar disso, não identificamos ferramentas voltadas para o apoio ao ensino das extensões espaciais de SQL, e constatamos que softwares desktop como QuantumGIS e gvSIG, embora consigam se comunicar com gerenciadores de bancos de dados geográficos, não possuem recursos que permitam fazer consultas em SQL e visualizar os resultados na forma de mapas na tela.

O presente artigo introduz GeoSQL, um ambiente online para o aprendizado de extensões espaciais de SQL, ferramenta computacional de apoio ao ensino em laboratório ou individualizado, pela Web, dos conceitos e funções que diferenciam bancos de dados geográficos dos convencionais. A Seção 2 apresenta trabalhos voltados ao ensino de SQL. A Seção 3 traz uma visão geral da funcionalidade e recursos do GeoSQL. A Seção 4 encerra o artigo, trazendo conclusões e listando trabalhos futuros.

## 2. Trabalhos Relacionados

Existem diversas ferramentas voltadas para o ensino da linguagem SQL, muitas delas disponíveis online. A literatura da área de ensino em computação descreve algumas iniciativas. Sadiq and Orłowska (2004) desenvolveram o SQLator, um ambiente online para aprendizado de SQL. O SQLator dispõe de um tutorial integrado que apresenta conceitos fundamentais, oferece diversos bancos de dados para prática, cada qual com um conjunto de questões de teste, e permite a execução real das consultas sobre os bancos de dados de prática. Seu principal diferencial é uma função que executa (usando o MS-Server) e verifica o resultado das expressões submetidas pelo estudante, juntamente com recursos para acompanhar o desempenho individual dos estudantes.

Aproximadamente a mesma funcionalidade do SQLator está disponível em LEARN-SQL (Abelló et al., 2008), que implementa uma arquitetura diferente, baseada em serviços Web. Aspectos de avaliação do desempenho de estudantes na formulação de consultas SQL são explorados por Prior (2003). Pereira and Resende (2012) apresentam uma avaliação ampla de ferramentas para ensino de bancos de dados, incluindo modelagem pelo modelo de entidades e relacionamentos (ER), álgebra relacional e SQL. É proposto um ambiente novo, projetado a partir do que foi observado nas ferramentas avaliadas.

Nenhum dos trabalhos analisados inclui o ensino de SQL espacialmente estendido. Até onde foi possível verificar, não existem ferramentas que disponham dessa capacidade, principalmente considerando que o resultado de consultas SQL espaciais é, frequentemente, de natureza geográfica e precisa ser apresentado graficamente. Além disso, a visualização de resultados muitas vezes só faz sentido se superposta a algum tipo de mapeamento básico, usado como *background* para prover contexto visual ao resultado de uma consulta. A seção seguinte apresenta nossa proposta para esse tipo de ambiente.

## 3. O GeoSQL

O GeoSQL<sup>1</sup> oferece uma interface na qual o usuário pode submeter uma consulta SQL a um banco de dados disponível previamente e obter as respostas na tela. Caso a resposta inclua algum atributo geográfico, uma visualização correspondente é produzida na

---

<sup>1</sup><http://geo.lbd.dcc.ufmg.br/geosql>

aba denominada *Mapa*. As saídas visuais de diversas consultas podem ser apresentadas simultaneamente, na usual metáfora de camadas. Naturalmente, é possível manipular a ordem de apresentação dessas camadas, e também definir as cores de apresentação dos objetos em cada camada. Com isso, mapas mais complexos podem ser produzidos passo a passo, e o resultado de consultas pode ser apresentado em contexto, *i.e.*, sobreposto a *background* contendo um mapeamento básico. Uma vez apresentados, os resultados das consultas podem ser explorados, usando recursos como *pan* e *zoom*. A parte textual do resultado da execução dos comandos é apresentada na aba chamada *Resultado*. Na aba *Tutorial*, como o próprio nome informa, o usuário pode encontrar um tutorial descrevendo o modo de uso da ferramenta. A interface do GeoSQL oferece a possibilidade de visualizar o esquema físico do banco de dados utilizado para as consultas, através da aba *Esquema*. Para viabilizar esse recurso, a estrutura das tabelas é capturada e armazenada. Um mecanismo foi implementado para atualizar a apresentação do esquema sempre que ocorrer alguma alteração em sua estrutura.

Inicialmente, o usuário indica o banco de dados com o qual deseja trabalhar. O administrador do GeoSQL pode adicionar vários bancos de dados ao ambiente. Para isso, é necessário criar um diretório no servidor, contendo um arquivo denominado *connection.php*, no qual são definidas constantes que indicam (1) o nome do plugin a utilizar para efetuar a conexão ao gerenciador de bancos de dados (SGBD), (2) o caminho e o nome do banco de dados a acessar, (3) o nome do usuário que acessará o banco e sua respectiva senha. A administração das permissões correspondentes a esse usuário padrão é realizada no próprio SGBD.

Os comandos SQL são digitados no campo de texto logo acima do botão *consultar*. Quando a consulta é disparada, uma função verifica se o comando é um SELECT. Caso seja, uma tabela temporária é criada para receber o resultado da consulta e nela se busca a primeira ocorrência de uma coluna geométrica. No caso específico do PostgreSQL, para verificar se uma coluna contém dados geográficos, basta verificar se seu tipo é igual a *geometry* ou *geography*. Uma vez obtido o nome da coluna geométrica, uma consulta do tipo **SELECT ST\_ASSVG(nome\_coluna\_geometrica) FROM nome\_tabela\_temporaria** é realizada. O resultado é um conjunto de dados geométricos codificados usando o *Scalable Vector Graphics* (SVG), padrão do World Wide Web Consortium (W3C) para a renderização de gráficos vetoriais. No SVG, sequências de pares de coordenadas são denominados *paths*, codificação geométrica que pode ser utilizada para a renderização dos pontos, linhas poligonais e polígonos utilizados na geometria de objetos geográficos. O resultado em formato SVG é, então, encaminhado para renderização do lado do cliente, diretamente na página HTML do GeoSQL, usando as tags adequadas. O resultado textual da consulta é produzido e encaminhado à aba *Resultado*, excluindo-se as colunas geométricas, cuja visualização textual não é de grande interesse para o usuário (Figura 1).

Após ser encaminhada ao servidor, cada consulta passa a compor um histórico que se encontra do lado direito da tela, logo abaixo dos dois ícones que se localizam no canto superior direito da aplicação. Essa lista permite que os resultados de consultas textualmente idênticas a outras já realizadas pelo usuário sejam refeitas com maior rapidez pois seus resultados são armazenados na sessão PHP de cada usuário até o término de sua conexão. Para se refazer a consulta, basta que o usuário clique sobre o ícone do lápis constante na caixa da consulta salva. Caso deseje, o usuário também pode excluir uma

pesquisa dessa lista, clicando sobre o ícone com um 'X' no canto inferior esquerdo de uma consulta armazenada. Ainda sobre cada consulta, os terceiro e quarto ícones permitem que o usuário modifique a cor da linha e o padrão de preenchimento dos *paths* de uma consulta com resultado plotado na aba *Mapa*.

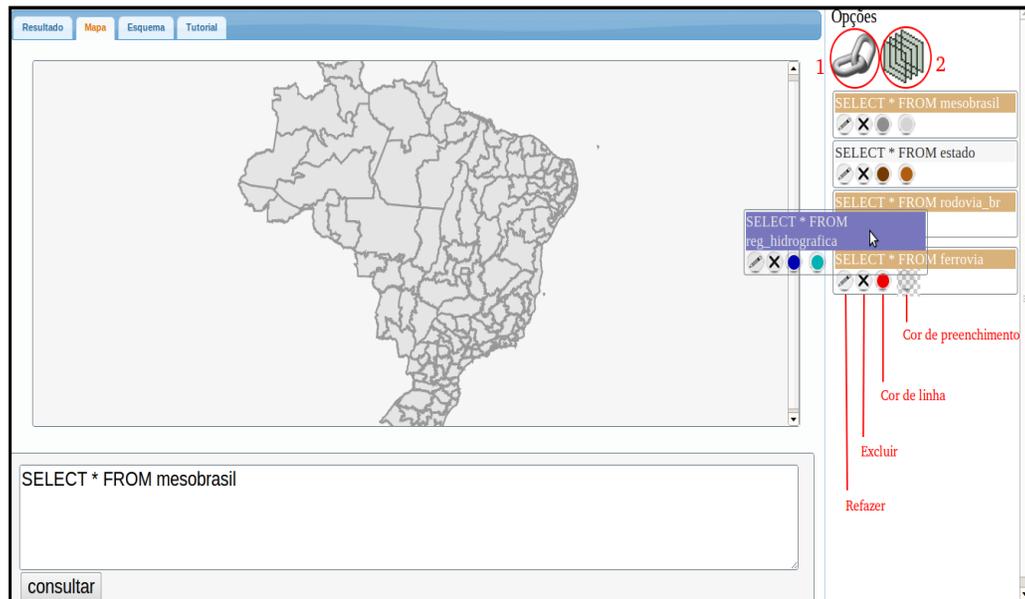
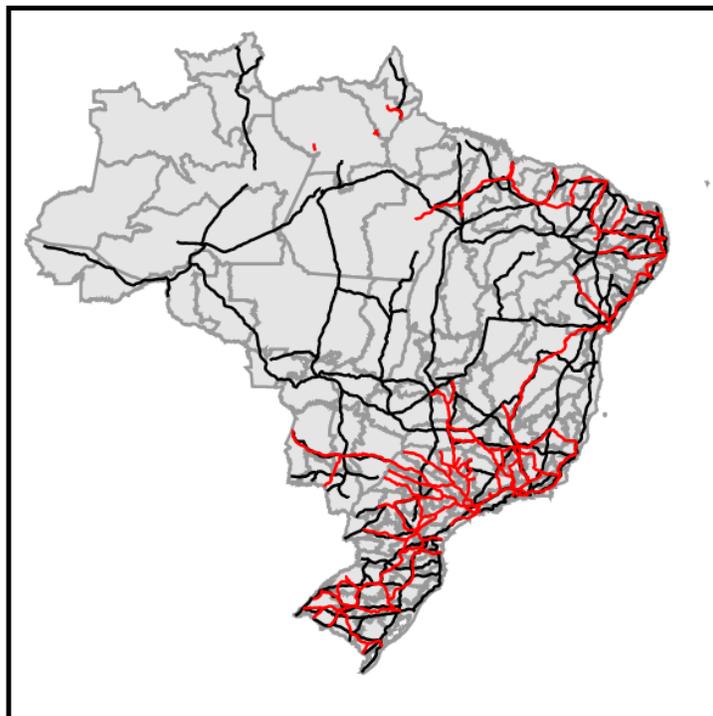


Figura 1. Aba *Mapa* em destaque e histórico de consultas

Dentre os dois ícones localizados no canto superior direito, o da esquerda (1) atua como um compactador do estado de todas as informações presentes no ambiente do usuário em determinado momento, gerando um *link* que pode ser acessado em momento posterior para recuperação de todos os dados, incluindo cores de mapas, tabelas e lista de consultas salvas, preservando-se sua ordem. Essa função é de grande utilidade para quando, por exemplo, um aluno deseja enviar o resultado de um conjunto de consultas a seu professor para avaliação em vez de ter que montar um arquivo específico com cada resposta. O ícone mais à direita (2) define uma operação de superposição entre as camadas de objetos SVG correspondentes às consultas previamente selecionadas. De cima para baixo na lista de consultas selecionadas (destacadas em laranja), a primeira consulta corresponderá à camada inferior da visualização, enquanto a última corresponderá à camada superior. Com o objetivo de permitir o reordenamento das diversas camadas, cada consulta da lista pode ser deslocada para cima ou para baixo em operações *drag and drop* sobre cada elemento (em roxo na figura).

Ainda na Figura 1 podemos visualizar cinco consultas já realizadas, com destaque para a consulta no topo da lista, a qual foi refeita com nova coloração de linha e de preenchimento. Observe-se que o mouse está deslocando uma consulta de sua posição original para um local mais abaixo na lista e existem três consultas selecionadas para posterior realização de *merging*. Na Figura 2 vemos o resultado da operação de sobreposição referente às três consultas selecionadas na Figura 1. Nele, consta ao fundo uma camada contendo os limites de mesorregiões brasileiras, seguido por uma segunda camada, com

*paths* delineados em preto, indicando as rodovias e, por fim, uma camada de linhas em vermelho, indicando todas as ferrovias brasileiras contidas no banco.



**Figura 2. Superposição de camadas originárias de consultas distintas**

Outras características interessantes do GeoSQL são o suporte à aplicação de *zoom in* e *zoom out* utilizando o botão de *scroll* do mouse e o deslocamento (*pan*) do mapa plotado na aba *Mapa* arrastando e soltando o objeto SVG apresentado. Como exemplo, mostramos que o objeto SVG plotado na aba *Mapa* apresentada na Figura 1 foi submetido às operações de *zoom in* e *pan* antes de ser retirado o *screenshot* da tela. Além disso, baseado no aprendizado das colunas presentes nas tabelas do banco, o campo de realização de consultas informa opções de *code completion* em uma lista que aparece dinamicamente abaixo da área de texto à medida que o usuário digita sua consulta.

Todo o código do GeoSQL foi desenvolvido em PHP e jQuery utilizando os plugins jQueryui, jPicker e svgPan. Também foi implantado suporte para comunicação com os SGBDs MySQL e PostgreSQL associado com a extensão espacial PostGIS, ambos sendo intermediados por um servidor Apache <sup>2</sup>.

#### **4. Conclusão e Trabalhos Futuros**

Apesar de ainda não se encontrar em estágio final, o uso do GeoSQL na prática demonstrou o potencial da ferramenta para o ensino (ou autoinstrução) sobre extensões espaciais

<sup>2</sup>Respectivamente <http://www.php.net>, <http://jquery.com/>, <http://jquery.com/ui>, <http://code.google.com/p/jpicker/i>, <http://code.google.com/p/svgpan/>, <http://www.mysql.com/>, <http://www.postgresql.org/>, <http://postgis.refractory.net/> e <http://www.apache.org>

da linguagem SQL. Nada impede, no entanto, que o GeoSQL seja utilizado também no ensino da linguagem SQL convencional. As decisões tecnológicas e de implementação da ferramenta foram tomadas de modo a simplificar seu gerenciamento e administração, permitindo criar ambientes para aprendizado básico (p.ex., com permissão apenas de consulta e ensino de comandos SELECT). O GeoSQL também não exige a instalação de quaisquer pacotes na máquina do cliente, pois pode ser operado utilizando apenas um navegador comum que consiga renderizar objetos SVG.

Uma extensão do nosso trabalho se destinará à análise da melhor maneira de realizar uma compressão mais eficiente dos dados que trafegam entre o servidor e o cliente. Pela estrutura modular e conteúdo textual das tags SVG, acreditamos que a utilização de compressão em nível de *paths* ou conjuntos unicamente identificados deles possa contribuir para a formação de um cache compartilhado entre clientes, de modo a diminuir consideravelmente a quantidade de dados processados a cada consulta e encaminhados a cada cliente. Outra evolução será estudada no sentido de implementar recursos de segurança que permitam a execução de comandos de manipulação de dados, tais como CREATE TABLE, ALTER TABLE ou UPDATE, dependendo de permissões previamente definidas no SGBD. Com isso, o GeoSQL poderia servir de interface para o aprendizado de todos os diferentes comandos de SQL.

Vislumbramos também a possibilidade de realizarmos modificações no código-fonte para adaptarmos os resultados das consultas para renderização dos dados geográficos em dispositivos móveis como *smartphones* e *tablets*, no intuito de tornar nossa aplicação ainda mais abrangente. Estão nos planos também outros tipos de extensões, estas de natureza didática, como a avaliação automática das consultas, o acompanhamento de turmas de alunos e o armazenamento de listas de exercícios.

## Referências

- Alberto Abelló, M. Elena Rodríguez, Toni Urpí, Xavier Burgués, M. José Casany, Carme Martín, and Carme Quer. LEARN-SQL: Automatic Assessment of SQL Based on IMS QTI Specification. *Advanced Learning Technologies, IEEE International Conference on*, 0:592–593, 2008.
- Max J. Egenhofer. “Why not SQL!”. *International journal of geographical information systems*, 6(2):71–85, 1992.
- George Percivall. OpenGIS Reference Model. *OpenGIS Reference Model, Open Geospatial Consortium, Inc*, 2003.
- Juliana Alves Pereira and Antônio Maria Pereira Resende. Uma análise dos ambientes de ensino de banco de dados. *Anais do VIII Simposio Brasileiro de Sistemas de Informação*, pages 755–766, 2012.
- Julia Coleman Prior. Online assessment of SQL query formulation skills. In *Proceedings of the fifth Australasian conference on Computing education - Volume 20, ACE '03*, pages 247–256, Darlinghurst, Australia, Australia, 2003. Australian Computer Society, Inc.
- Shazia Sadiq and Maria Orlowska. SQLator: An Online SQL Learning Workbench. In *In Proceedings of the 9th annual SIGCSE conference on Innovation and technology in computer science education ITiCSE '04*, page 30, 2004.