



Geoinformatika

VII – Analýza dat jaro 2014

Petr Kubíček

kubicek@geogr.muni.cz

**Laboratory on Geoinformatics and Cartography (LGC)
Institute of Geography
Masaryk University
Czech Republic**

- **Prostorové analytické možnosti GIS tvoří jádro GIS, tedy to, co GIS odlišuje od ostatních informačních systémů.**
 - **Mezi otázky, na které nám GIS umožňuje odpovědět patří:**
 - Co se nachází na?
 - Kde se nachází?
 - Jaký je počet?
 - Co se změnilo od?
 - Co je příčinou?
 - Co když?
- WHAT? – WHERE? – WHEN?**



Analytické nástroje GIS

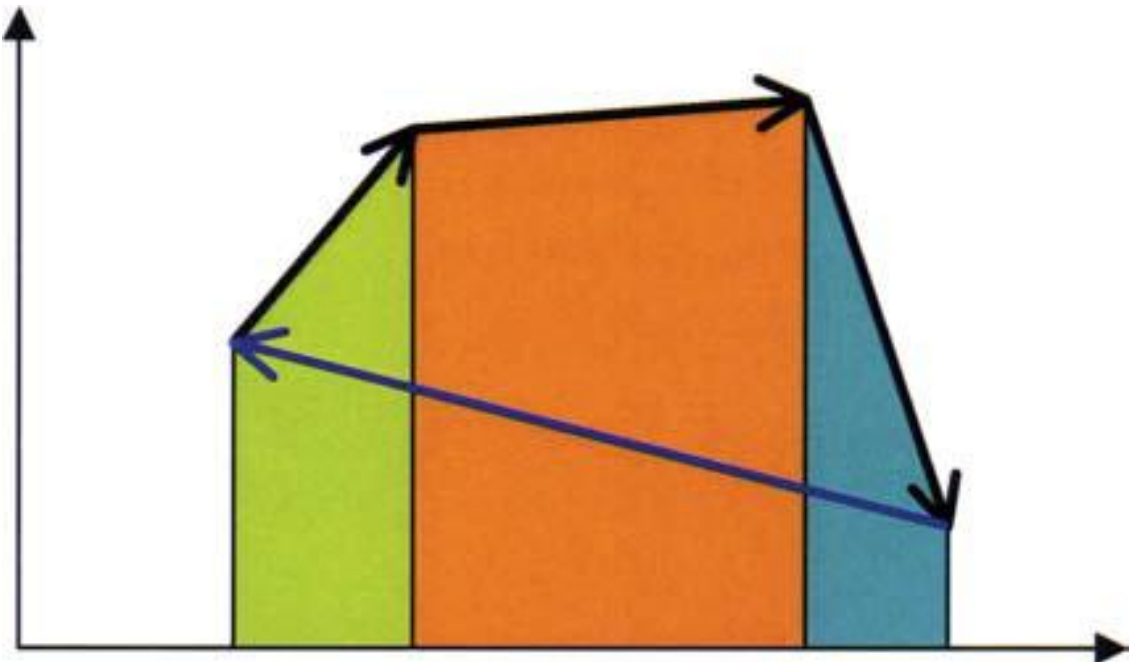
Analytické možnosti GIS můžeme rozdělit do následujících skupin:


- měřicí funkce,
- **atributové i prostorové dotazy**(nástroje na prohledávání databáze),
- topologické překrytí,
- mapová algebra,
- vzdálenostní analýzy,
- analýzy sítí,
- analýzy modelu reliéfu a dalších povrchů,
- statistické analýzy.

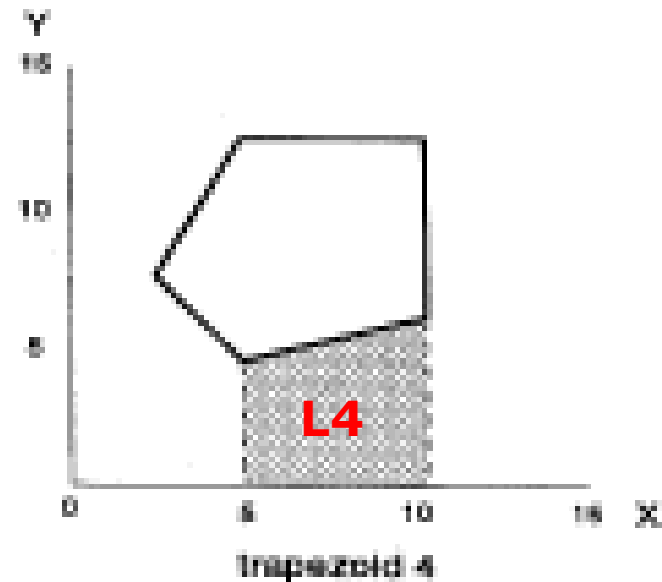
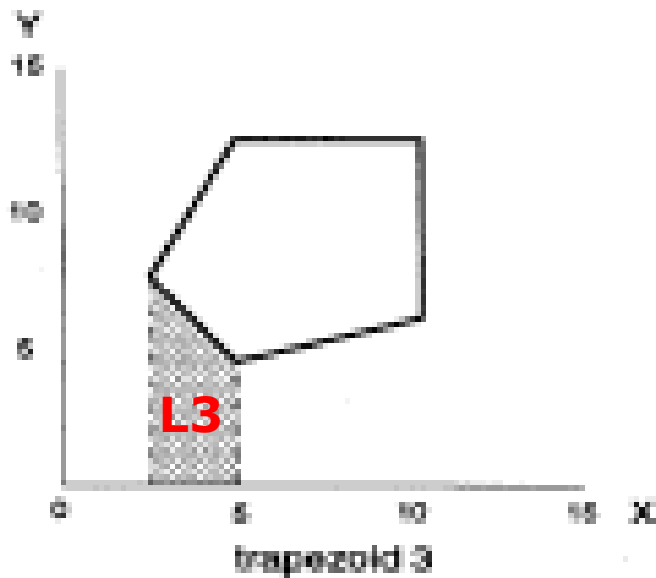
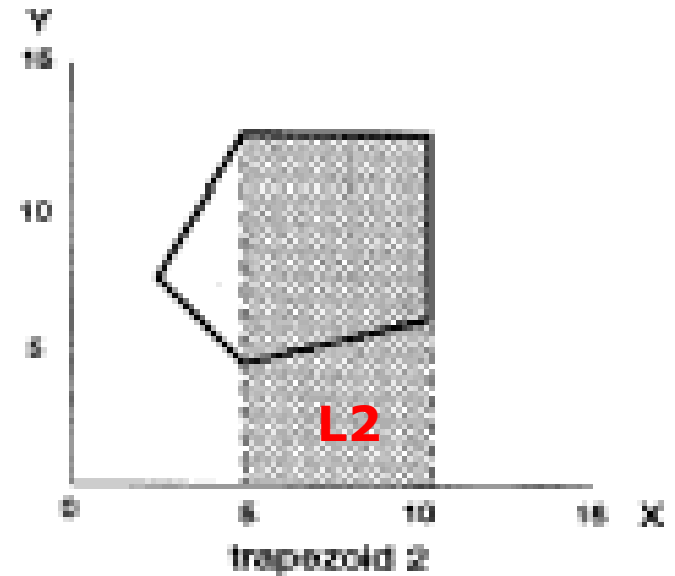
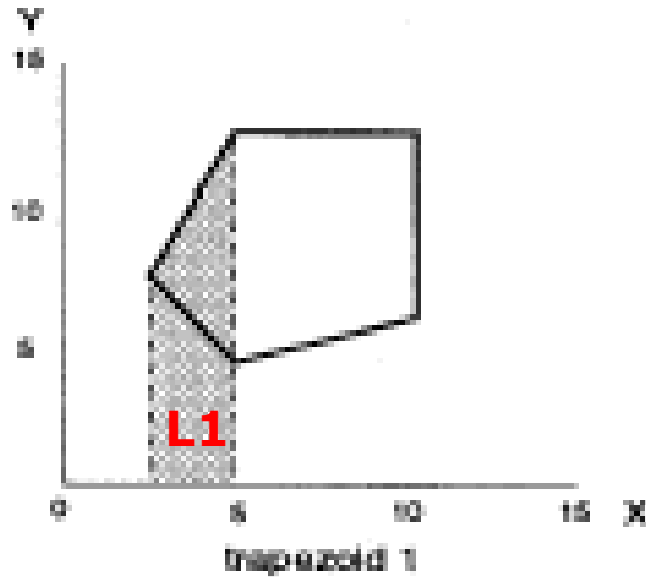
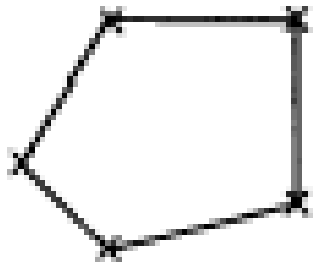
- GIS poskytují funkce na **měření vzdáleností a ploch**.
- Geografické informační systémy umí používat **různé délkové jednotky** (stopy, cm, ...), případně mezi nimi automaticky provádět převody.
- Při projektech v malém měřítku, a tedy většího plošného obsahu, má na měření vliv také **zakřivení zemského povrchu**, takže GIS produkty mohou **umožňovat započítat i tento faktor**.
 - konformní - nedochází ke zkreslení úhlů,
 - ekvivalentní - nedochází ke zkreslení ploch,
 - ekvidistantní - nezkrsluje délky ve směru určité soustavy křivek.
 - kompenzační - dochází k deformaci všech geometrických prvků (úhlů, délek i ploch), ale hodnoty deformací nejsou extrémní.

Měření ploch

- **Ruční měření – čtvercová metoda, planimetrie, ... časově náročné, nepřesné (opakovaná měření).**
- **Plochy**

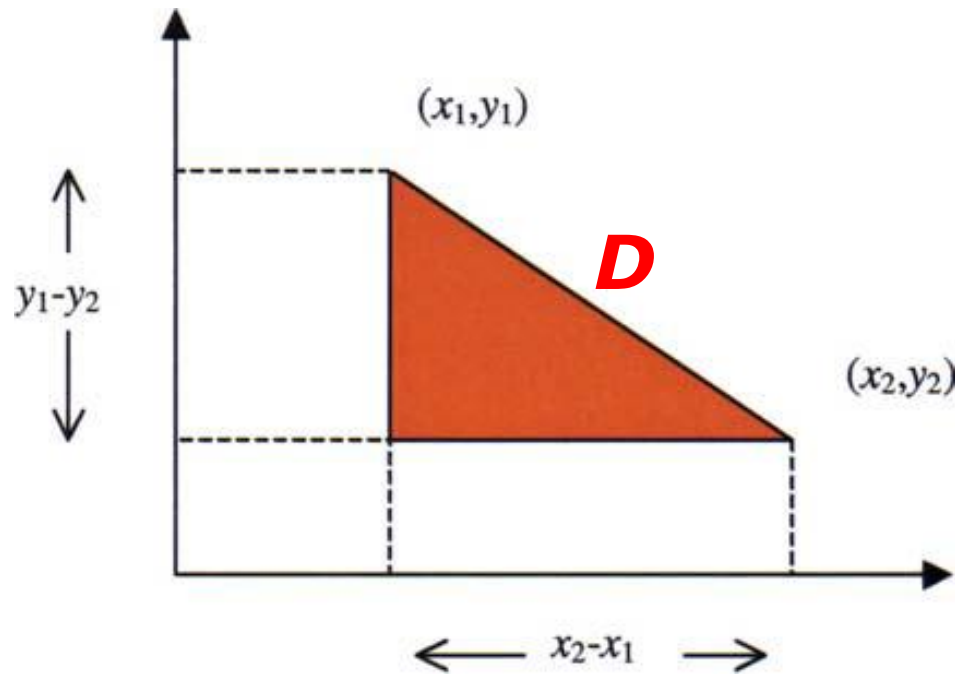



$$\text{Celková plocha} = (L1 + L2) - (L3 + L4)$$



Měření vzdáleností

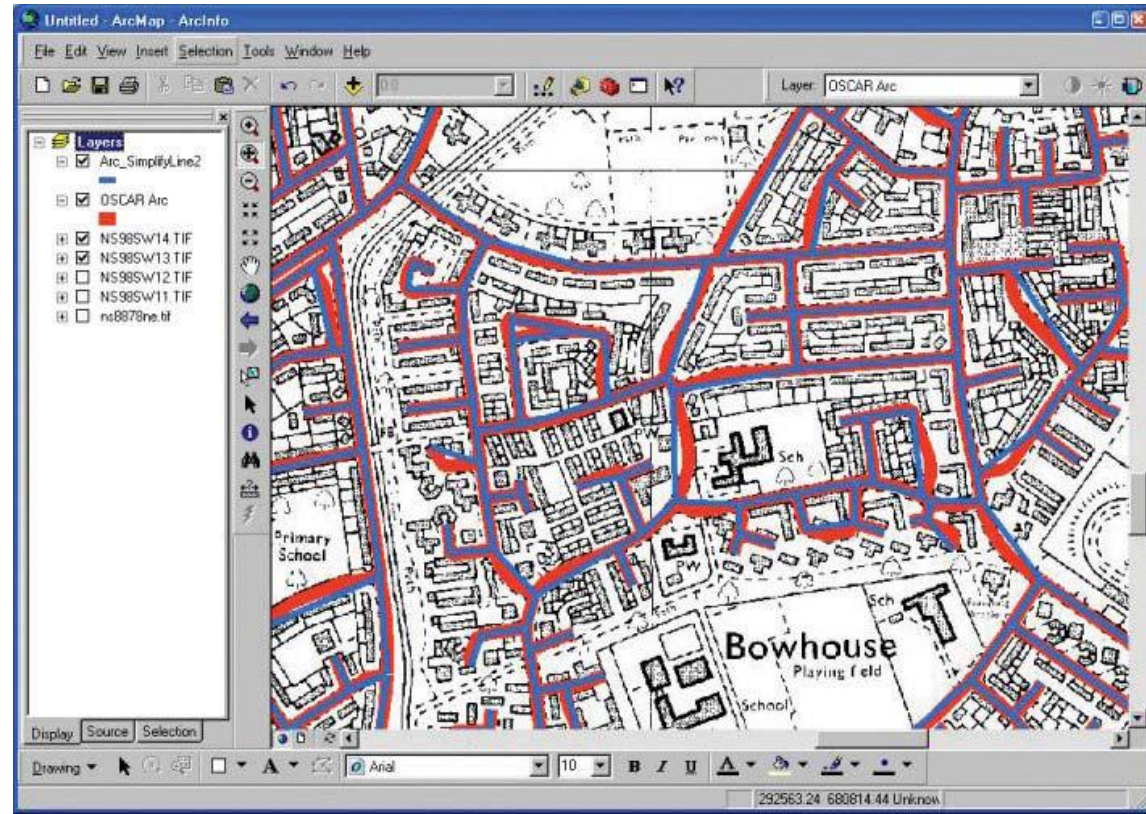
- Metrika – nejkratší vzdálenost v rovině.



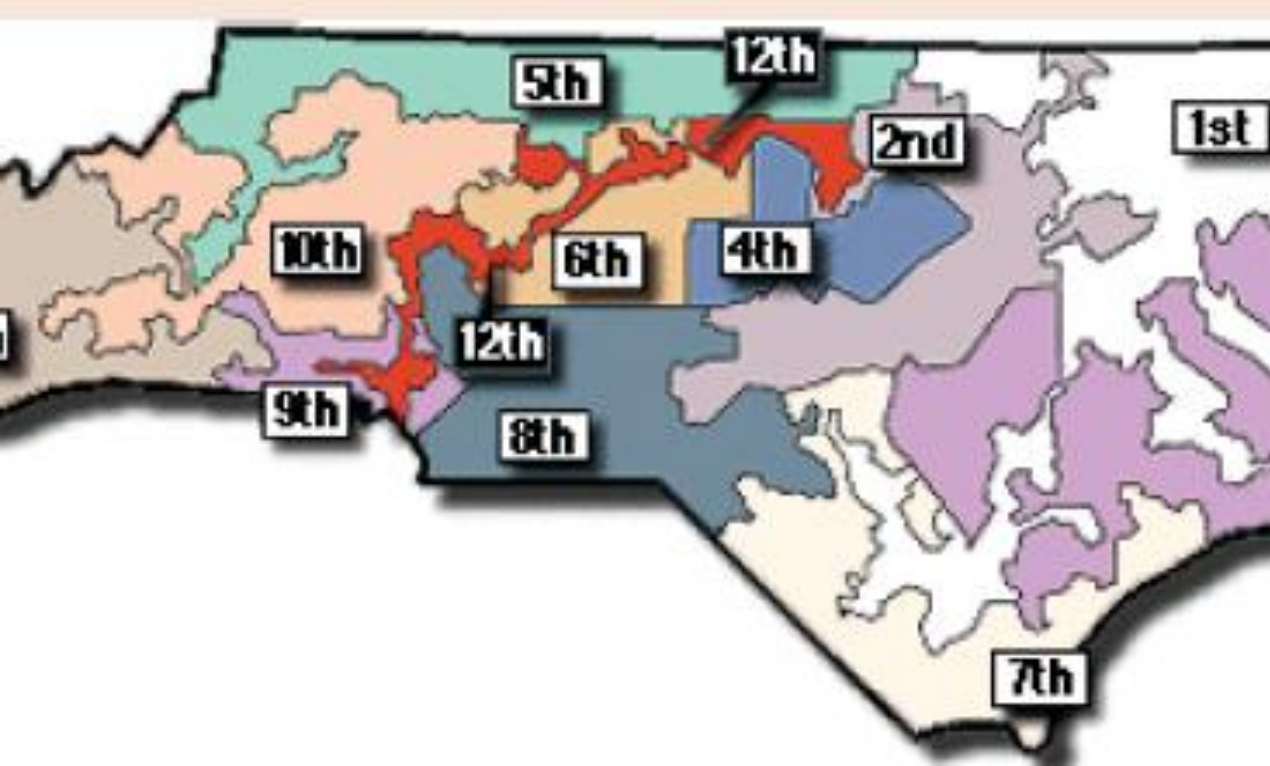
$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Komplexní měření vzdálenosti

- Zjednodušená reprezentace objektivní reality.
- „A GIS will almost always underestimate the true length of a geographic line.“



Měření tvaru



- S - kompaktnost
- P - obvod (perimeter)
- A - plocha (area)
- $3,54 = 2x\sqrt{\pi}$
- S=1 pro kruh.

$$S = P / 3.54\sqrt{A}$$



Dotazy na geografická data

- Dotazováním se vybírají údaje, které odpovídají specifickému kritériu nebo podmínce.
- Dotazovací operace má obvykle tři hlavní komponenty:
 - 1) Specifikace údajů, kterých se týká.
 - 2) Formulace podmínek, kterým musí údaje vyhovovat.
 - 3) Instrukci, co se má na vybraných údajích vykonat.
- Dotaz (Query) má tedy obecně následující strukturu:
„vyber z údajů typu T takové, které vyhovují podmínce P a vykonej na nich operaci O.“

Typy dotazů

Dotazy můžeme v GIS rozdělit na:

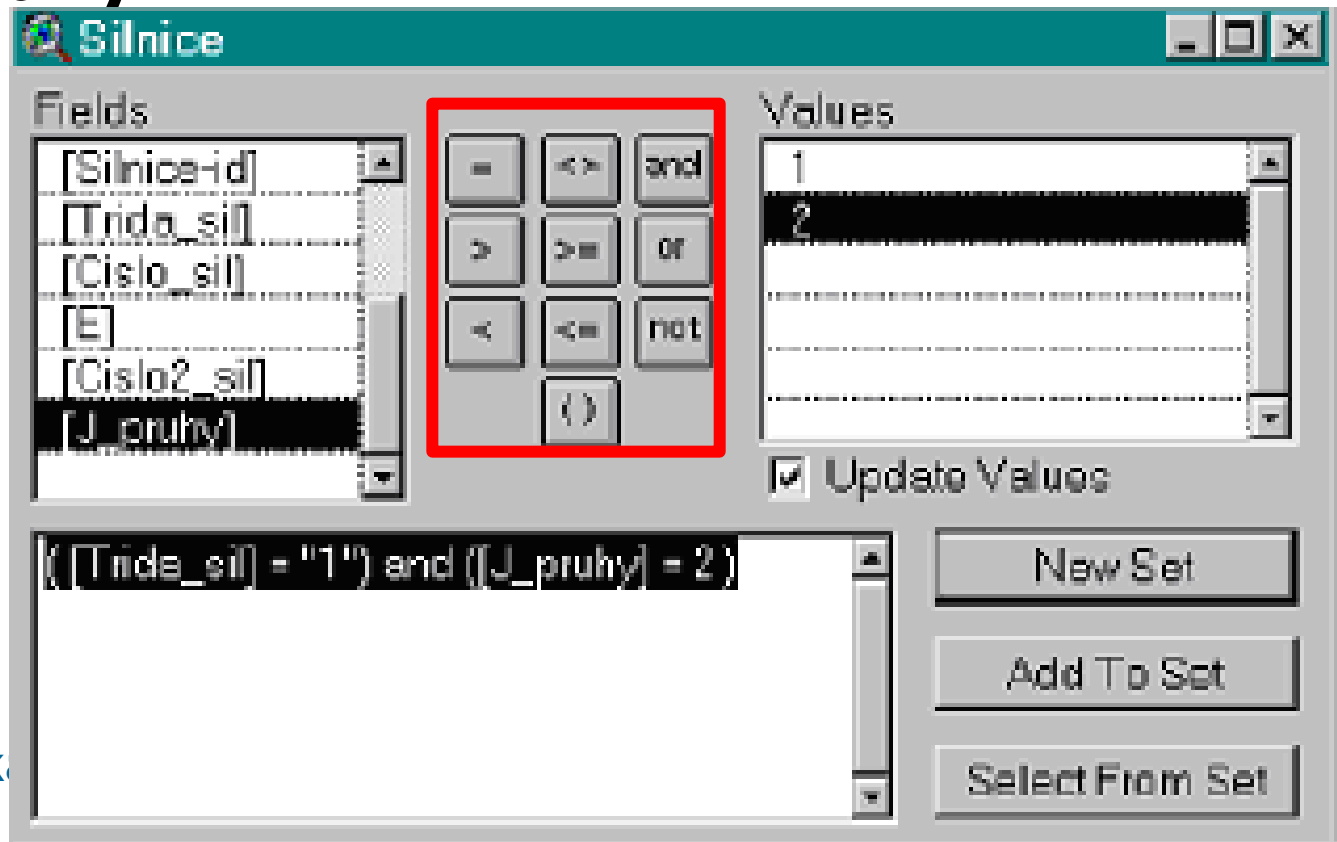
- **Atributové** - dotaz typu: "které geografické objekty (lokality) mají definovanou vlastnost".
 - *Například: "Zvýrazni všechna města v ČR, která mají více jak 10 000 obyvatel".*
- **Prostorové** - dotaz typu: "co se nachází na tomto místě, co se nachází v této oblasti".
 - *Například: "Zvýrazni všechna města v ČR, která leží v Jihomoravském kraji".*
- **Kombinované** - dotaz typu: "které objekty splňují definovanou vlastnost a zároveň se nacházejí v nějaké oblasti"
 - *Například: "Zvýrazni všechna města v ČR, která mají více jak 10 000 obyvatel a zároveň leží v Jihomoravském kraji".*

Atributové dotazy

Dotazují se na **atributy (vlastnosti)** geografických dat!

- **Lze je uskutečnit různými způsoby:**
 - **Identifikace** jednotlivého **objektu na základě** jeho jména, označení či jiného **atributu**.
 - *Př. Vypiš všechny vlastnosti dálnice D5 ve vrstvě silnic.*
 - **Vyhledání** všech **objektů splňující intervalové či logické podmínky** jednoho nebo více atributů.
 - *Př. Vyber všechny silnice 1. třídy mající dva jízdní pruhy.*

- **Vrstva silnic má dva atributy:**
 - třídu silnice (1,2,D,R,o) a počet pruhů (1,2).
- **Vyber všechny silnice 1. třídy, které mají 2 jízdní pruhy**



Silnice

Fields

- [Silnice-id]
- [Trida_sil]
- [Cislo_sil]
- [E]
- [Cislo2_sil]
- [J_pruhy]

Values

- 1
- 2

Update Values

Logical Operators:

- =
- <>
- and
- >
- >=
- or
- <
- <=
- not
- ()

Query: ([Trida_sil] = '1') and ([J_pruhy] = 2)

Buttons: New Set, Add To Set, Select From Set

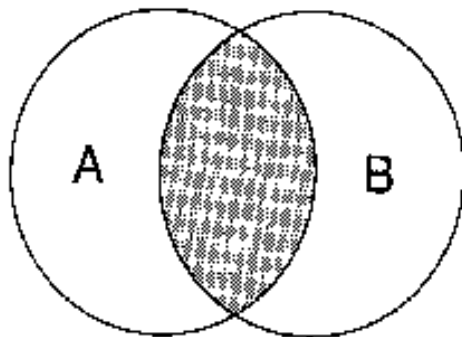


Dotazy na geografická data

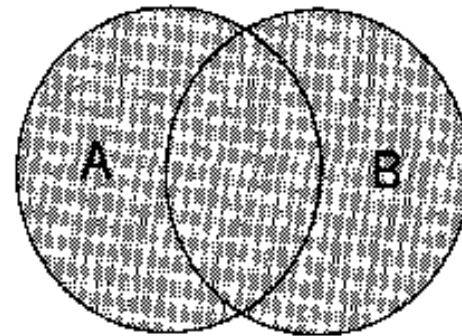
- Je vidět, že pro vyhledávání **intervalových podmínek** je možné použít operátorů $<, >, =, <=, >=, <>$.
- Intervalové podmínky jdou dále kombinovat pomocí **logických operátorů** (AND, OR, NOT) využívajících pravidel Booleovské logiky.
- Dotaz (Query):
- **SELECT * FROM SILNICE WHERE TRIDA_SIL="1" and J_PRUHY=2**

Logické operátory

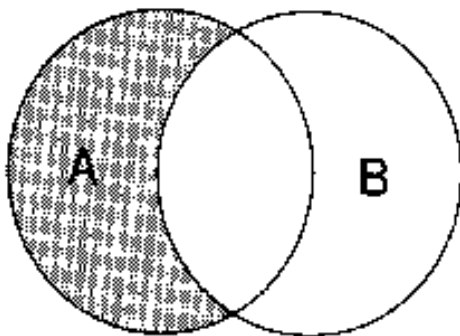
- Pro dva prvky.



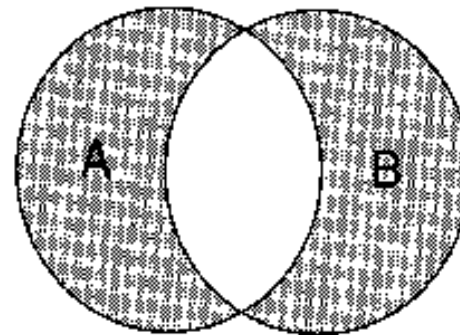
A AND B



A OR B



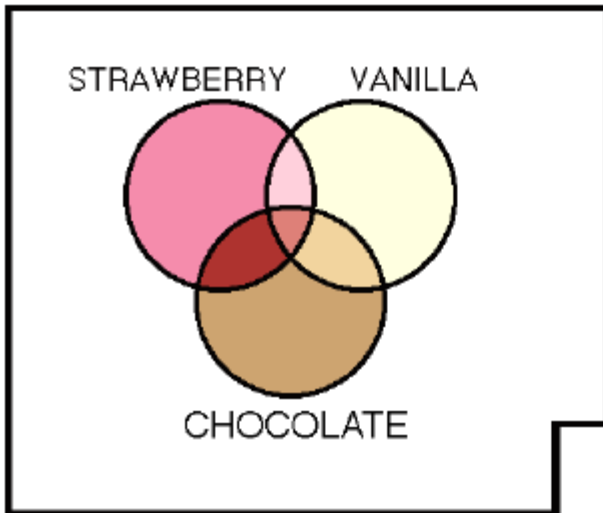
A NOT B



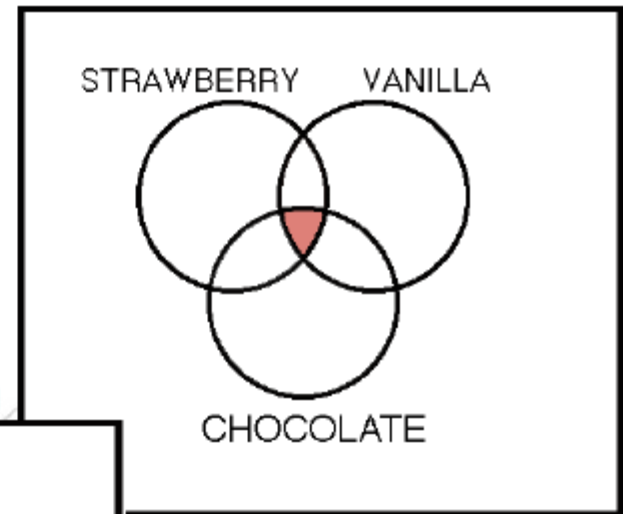
A XOR B

Logické operátory - kombinace

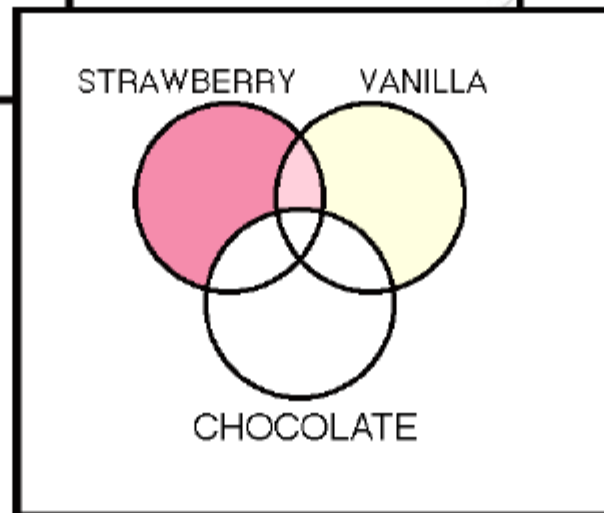
S or V or Ch



S and V and Ch

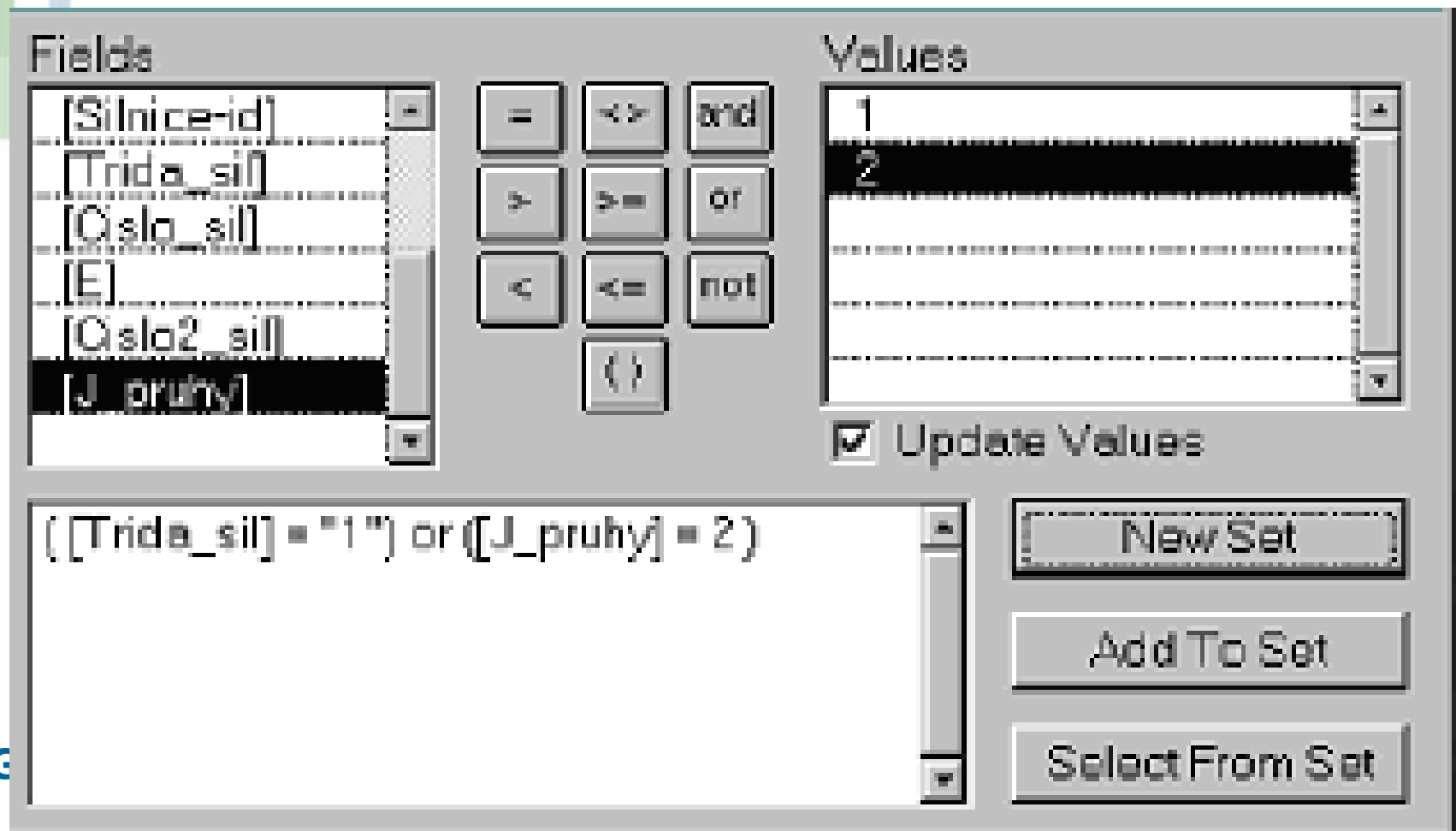


S or V and not Ch



Příklad

- Vyber všechny silnice které jsou první třídy nebo mají dva jízdni pruhy.



The screenshot shows a query builder interface with the following components:

- Fields:** A list of fields including [Silnice-id], [Trida_sil], [Oslo_sil], [E], [Oslo2_sil], and [J_pruhy]. The field [J_pruhy] is currently selected.
- Operators:** A set of buttons for logical operations: =, <>, and, >, >=, or, <, <=, not, and parentheses ().
- Values:** A list of values with '1' and '2' visible. The value '2' is highlighted.
- Update Values:** A checked checkbox labeled 'Update Values'.
- Expression:** A text area containing the logical expression: `([Trida_sil] = "1") or ([J_pruhy] = 2)`.
- Buttons:** Three buttons on the right: 'New Set', 'Add To Set', and 'Select From Set'.

- **Vyber všechny silnice s dvěma jízdními pruhy, které nejsou první třídy.**

Silnice

Fields

- [Silnice-id]
- [Trida_sil]
- [Cislo_sil]
- [E]
- [Cislo2_sil]
- [J_pruhy]

Values

- 1
- 2

Logical Operators: =, <>, and, >, >=, or, <, <=, not, ()

Update Values

Query: `([J_pruhy] = 2) and not ([Trida_sil] = '1')`

Buttons: New Set, Add To Set, Select From Set



Prostorové dotazy

- Dotazují se na **prostorové vlastnosti a vztahy (geometrii a topologii)** geografických dat!
- **Lze je uskutečnit různými způsoby:**
 - Identifikace geografického objektu na základě jeho **souřadnic**, a to buď ručně (zadáním souřadnic) nebo interaktivně (ukázáním na objekt myší).
 - **Prohledávání prostoru různých geometrických tvarů** (obdélníky, kružnice, polygony, linie) za účelem nalezení prvků, které splňují podmínku dotazu.



Prohledávání prostoru různých geometrických tvarů

- **Nejčastější podmínky prostorového dotazu:**
 - překrývají se – průnik,
 - dotýkají se (linií, bodem),
 - jsou obsaženy v nějaké oblasti/prvku,
 - obsahují nějaký prvek,
 - jsou identické,
 - jsou v nějaké vzdálenosti od určitého prvku/oblasti,
- ...

Prostorové predikáty

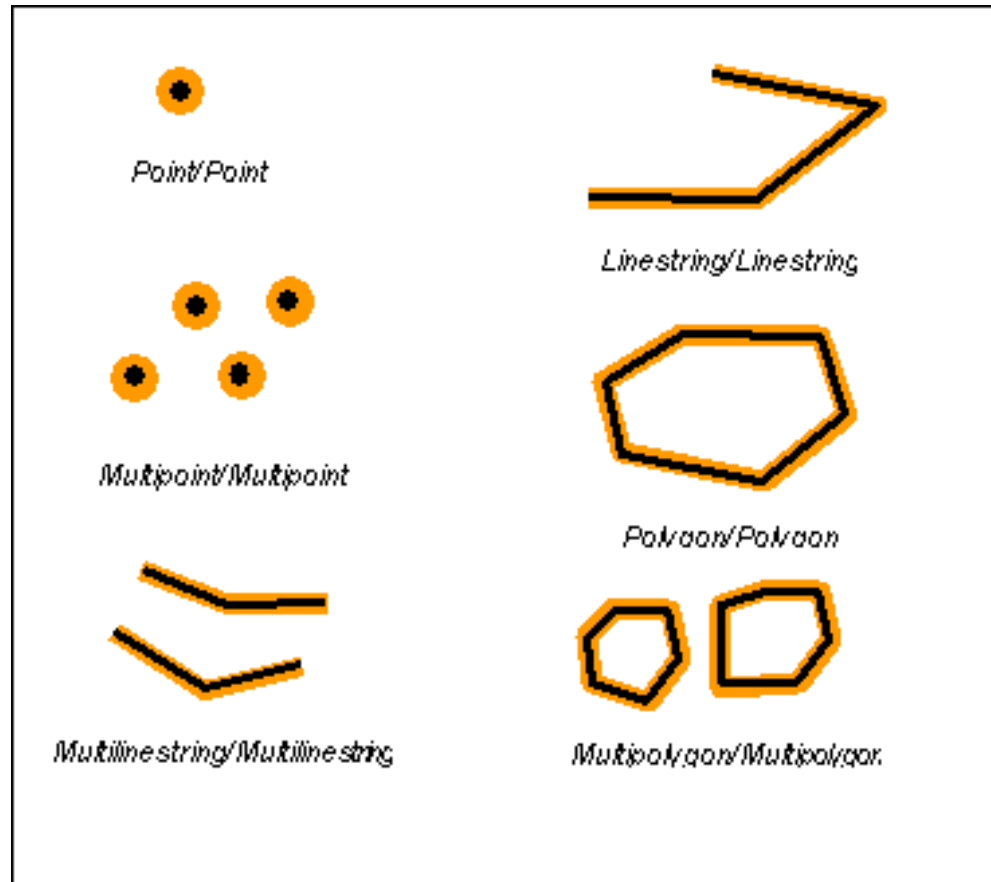
Boolean funkce vyjadřující specifické prostorové vztahy pro dvojici geometrických prvků. Pokud splňují podmínku = TRUE

Prvky mohou mít rozdílné geometrie (bod, plocha, linie...). Pouze X,Y souřadnice!

Predikáty zkoumají vždy vnitřek, hranice a vnějšek geometrického prvku.

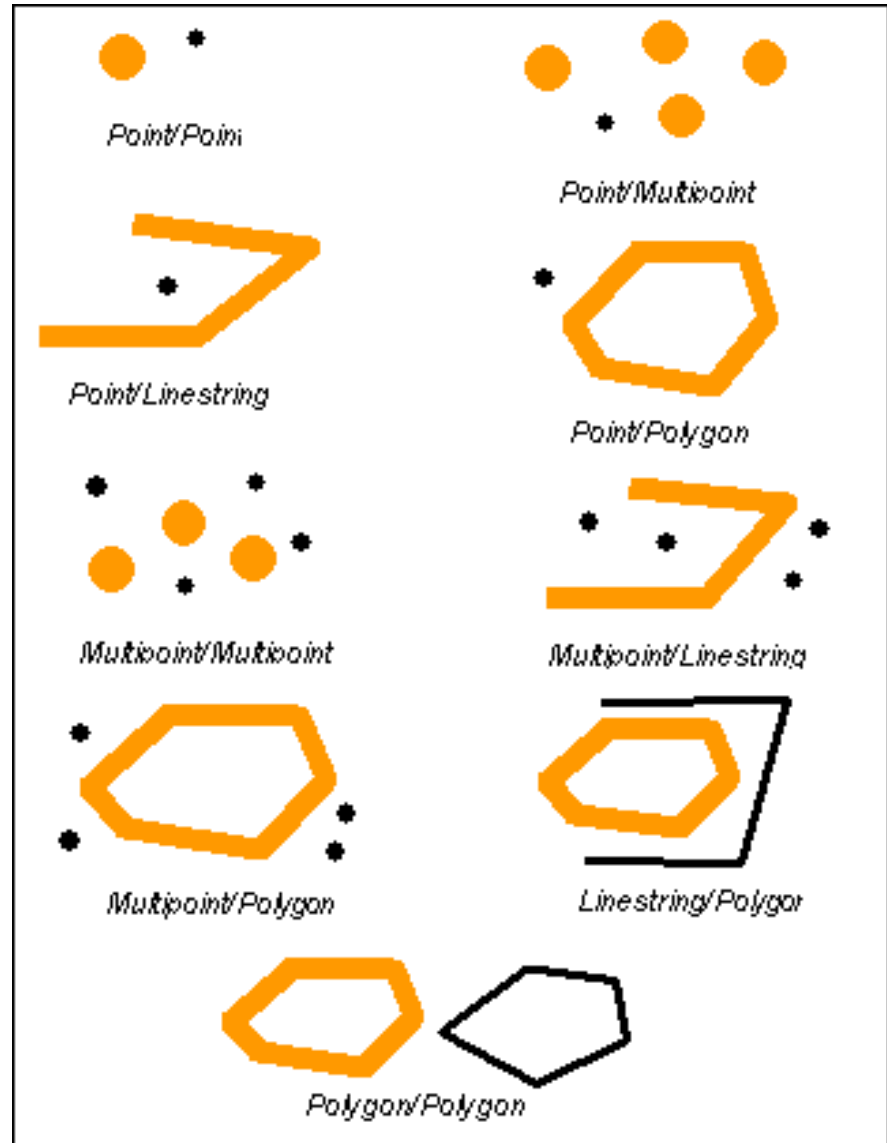
- **Equal**
- **Disjoint**
- **Intersects**
- **Touch**
- **Overlap**
- **Cross**
- **Within**
- **Contains**

- **Equal** returns **t (TRUE)** if **two geometries of the same type** have identical X,Y coordinate values.



Intersect and Disjoint

- Intersects returns t (TRUE) if the intersection **does not result in an empty set**.
- Disjoint returns t (TRUE) if the intersection of the two geometries **is an empty set**.



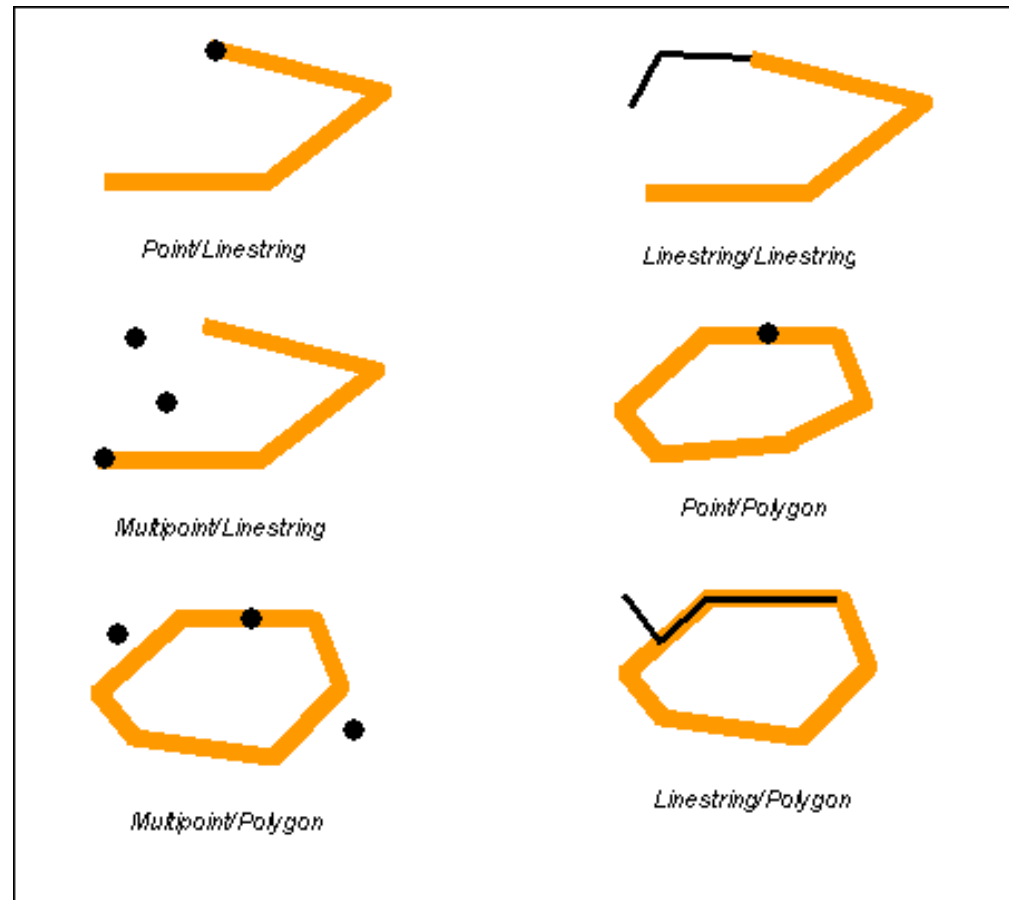
Touch

- **Touch returns t (TRUE) if none of the points common to both geometries intersect the interiors of both geometries.**

		b		
		Interior	Boundary	Exterior
a	Interior	F	T	*
	Boundary	*	*	*
	Exterior	*	*	*

		b		
		Interior	Boundary	Exterior
a	Interior	F	*	*
	Boundary	T	*	*
	Exterior	*	*	*

		b		
		Interior	Boundary	Exterior
a	Interior	F	*	*
	Boundary	*	T	*
	Exterior	*	*	*



Overlap

- **Overlap** compares **two geometries of the same dimension** and returns **t (TRUE)** if their intersection set results in a geometry different from both but of the same dimension.

Výsledek musí být tvarově odlišný, ale stejné geometrie.



Linestring/Linestring



Polygon/Polygon



Multipoint/Multipoint

Cross returns **t** (TRUE) if the intersection results in a geometry whose **dimension is one less than the maximum dimension of the two source geometries** and the intersection set **is interior to both source geometries**. **Cross** returns **t** (TRUE) for only multipoint/polygon, multipoint/linestring, linestring/linestring, linestring/polygon, and linestring/multipolygon comparisons.



Multipoint/Linestring



Multipoint/Polygon



Linestring/Linestring



Linestring/Polygon

- **Within** returns **t** (**TRUE**) if the first geometry is completely within the second geometry. **Within** tests for the exact **opposite** result of **contains**.

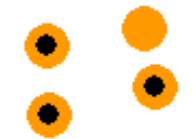
Within



Point/Multipoint



Point/Linestring



Multipoint/Multipoint



Multipoint/Linestring



Point/Polygon



Multipoint/Polygon



Linestring/Linestring



Linestring/Polygon

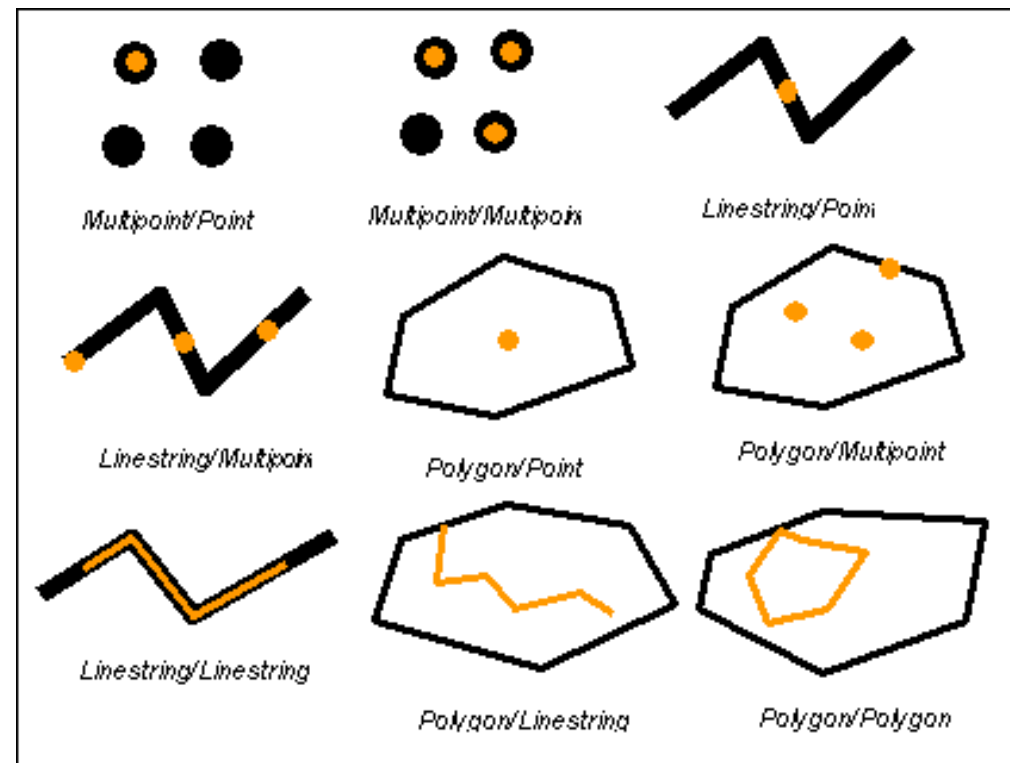


Linestring/Polygon

Contains

- Contains returns t (TRUE) if the second geometry is completely contained by the first geometry. The contains predicate returns the exact **opposite** result of the **within** predicate.**

		b		
		Interior	Boundary	Exterior
a	Interior	T	*	*
	Boundary	*	*	*
	Exterior	F	F	*



Kombinované dotazy

- **Jedná se v podstatě o řetězení a kombinování atributových a prostorových dotazů tak aby výsledek odpovídal požadované informaci.**
- **Jak atributové, tak prostorové dotazy pracují pouze s jednou informační vrstvou.**
- **Kombinované dotazy umožňují práci i s více vrstvami (či množinami objektů),**
 - Je zde opět možnost propojovat je pomocí operátorů Booleovské logiky, podobně jako u atributových dotazů.
 - Kombinované dotazy také zčásti mohou využívat topologické překrývání vrstev.

Dotazy vektor x rastr

- U **vektorové reprezentace** se zpracovávají údaje **atributových tabulek** připojených k jednotlivým vektorovým objektům.
- U **rastrové** se zpravidla zpracovávají údaje uložené v **buňkách** jednotlivých vrstev (teoreticky lze i u rastrových reprezentací mít připojené atributové tabulky).
- U **vektorů** je vždy vybrán **celý objekt**, u **rastrů** je vybírána vždy konkrétní **buňka** či skupina buněk.