Approval: 16th Senate Meeting

Course Number	**:** CS544
Course Name	**:** Formal Concept Analysis: Theory and Practice
Credits	**:** 3-0-0-3 **(**L-T-P-C)
Prerequisites	**:** CS208 – Mathematical Foundations of Computer Science or equivalent IC250 – Programming and Data Structure Practicum or equivalent
Intended for	**:** B.Tech. (EE. & CSE) /MS/MTech/PhD
Distribution	**:** Elective for B.Tech. III/IV year, MS, M.Tech., Ph.D**.**
Semester	**:** Even/Odd

1. Preamble:

Formal concept analysis emerged around 1980 as a mathematical theory of concepts and concept hierarchies. A concept is determined by its `extent' and its `intent': the extent consist of all objects belonging to the concept, while the intent contains all attributes shared by those objects. There exists a natural hierarchy between concepts, for example, the concept `human being' is a sub-concept of the concept `being' since every human being is also a being. FCA offers a structural and less numeric approach to data unlike many of today's data analysis methods that rely on numerical values. They often unfold the data, making it better accessible to the analyst's judgments and decisions. FCA finds applications in several areas of data and web mining, software engineering, artificial intelligence, etc. This course will feature some lectures from industry experts on how they apply FCA. This course will also include assignments where the students implement different FCA algorithms, and experiment with different real-world datasets. FCA research is still in its infancy; one aspect of this course is to highlight the open issues in FCA through survey of state-of-the-art research papers.

2. Course Modules with Quantitative lecture hours:

Formal contexts, formal concepts and concept lattice:

Formal context; conceptual scaling of many valued contexts; concept forming operators; basic mathematical structures behind FCA: Galois connections and closure operators; algebra of concepts - partial order and lattices; basic theorem; concept lattice diagrams. (8 hrs)

Centralized and distributed algorithms for concept lattice construction:

Clarification and reduction of formal concepts; basic algorithms for computing concept lattice – next closure algorithm, linear time closed item-set miner; distributed implementation of the algorithms for concept lattice construction – naive approach, map-reduce based close-by-one, distributed closed item-set miner, trade-offs between breadth-first search and depth-first search based approaches; concept interestingness measures – support, cue-validity, stability, lift and separation; incremental algorithms for constructing concept lattice. (12 hrs)

The canonical basis:

Attribute implications, computing closure under implications, learning implications with membership and equivalence queries; pseudo-closed sets and canonical basis of implications, finding pseudo-closed sets; attribute exploration algorithm and its variations; concept exploration with partial lattice. (10 hrs)

Applications of FCA:

Case-studies of concept exploration – exploring faulty data, exploration in a fuzzy setting, triadic data; analysis of social networks – social networks as formal contexts, individuality of social networks,

knowledge communities; FCA in databases: learning functional dependencies, extensions to knowledge graphs; neural network architecture based on concept lattice. (12 hrs)

3. Text book:

- Bernhard-Ganter and Sergei Obeidkov, "Conceptual Exploration", Springer 2016.
- Rokia Missaoui, Sergei O. Kuznetsov and Sergei Obiedkov, "Formal Concept Analysis of Social Networks", Lecture Notes in Social Networks, Springer 2017.

4. Reference Books:

- Research papers on FCA
- Open source softwares:
  - https://github.com/fcatools/
  - http://www.upriss.org.uk/fca/fcasoftware.html
- Radim Belhohlavek, "Introduction to Formal Concept Analysis", Olomouc 2008.
- Bernhard Ganter and Rudolf Wille, "Formal Concept Analysis: Mathematical Foundations", Springer 1999.

5. Similarity Content Declaration with Existing Courses:

S.N	Course Code	Similarity Content	Approx % of content
1	CS677 Knowledge representation and reasoning	Formal Concept Analysis	<10%

6. Justification for new course proposal if cumulative similarity content is > 30%: N.A.