Artificial Intelligence
CSE 5290/4301, Fall 2023

Instructor: Debasis Mitra, Ph.D.
Office: OEC 350 E-mail: dmitra 'at' cs.fit. edu
Class Home Page: http://www.cs.fit.edu/~dmitra/ArtInt/
Office Hours: MW 2-4pm, preferably, let me know first by e-mail (or by appointment)
Grading plan (Tentative, slight modifications may follow in weights):
Graduate stds: Takehomes/Quizzes: 20%, Exams 50%, Project: 30%
Undergraduate stds: Takehomes/Quizzes: 20%, Coding exercises/Project: 30%, Exams 50%

`Class 3:30 pm - 4:45 pm TR Olin Life Sciences 129`
`Class Zoom URL:` https://fit.zoom.us/j/97680524787
Access to this Table*: cs.fit.edu/~dmitra/ArtInt then, Spring 2023*
then, Continuously updated day-planner for Spring 2023

SYLLABUS AS ON THE FINAL EXAM OF **Fall 2022**
Background materials may be always questioned, they are pre-requsite.
Materials under *---* are excluded. 3.3-3.4.5 means 3.3 through 3.4.5
Observe the url, textbook pdf slides has /TextSlides/ and my ppt slides has /lectures/ in the url.

SEARCH
Informed Search: TextSlides Ch 4a, 4b, up to IDA* search and SMA*
Constraints: TextSlides Ch 5, Up to materials on my slide Constrained Optimization
Adversarial Search: TextSlides Ch6 up to Sl24-deterministic games
LOGIC
Propositional Logic: TextSlides Ch 7
FOL: TextSlides Ch 8
Inferencing in FOL: TextSlides Ch 9
PROBABILISTIC REASONING
Prior and Conditional probabilities, Probabilistic knowledge-base and inferencing: TextSlides Ch 13
Bayesian network: TextSlides Ch 14a
~~*Inferencing with Bayesian networ: TextSlides Ch 14b**~~
MACHINE LEARNING
Inductive learning, Decision Tree learning: TextSlides Ch 18
(In addition, follow my lecture slides for Machine Learning, all slides)
The following section are from textbook:
Evaluation 18.4-Model Selection 18.4.1
Regularization 18.4.3
Learning Theory 18.5.0 only
Regression 18.6 – 18.6.2
Classification 18.6.3 – 18.6.4
Neural Network 18.7 – 18.7.4 (exclude exotic varieties of NN on my slides)
Non-parametric models 18.8 – 18.8.4
SVM basics 18.9
Clustering basics (from my slides)
ETHICS

My slides in https://cs.fit.edu/~dmitra/ArtInt/lectures/AIethics.ppt

Florida Tech Academic Calendar: https://www.fit.edu/registrar/academic-calendar/spring-2021/

Detailed activities in Fall 2022 mapped on to the dates of Spring 2023, i.e., dates on the leftmost column should be correct but the activities in other two columns get continuously updated.

This table acts as my curricular map for the semester. It helps me to track progress and plan the next few days of the class, assignments, etc. I edit this frequently but sometimes it gets out of sync. Use it at your discretion.

Meetings: ~ 28. Lectures: Search 4, Constraints 3, Automated Reasoning 4, Probabilistic Reasoning 4, Machine Learning 5, Ethics 1,
Exams 3/4, Std presentations 2/3

This col is updated for Spr 2023	**Activities (Fall2022):**	**Comments**
Jan 10 T	AI an Introduction, and Get-to-know Pre-quiz (0 point but must submit).	**Pre-quiz is not online, not graded** *Zoom did not work!*
Jan 12 R	Required CS background list *Big O-notation (Algo-Intro slide 1-10),* *Djikstra algorithm (Algo-Graph slide 23-25),* ~~*Min spanning tree algorithm (Algo-Graph slide 54, 57-8)*~~ https://fit.zoom.us/rec/share/alXPlyk3RT6Kzv9htyeOpNvTC2LxeKqFnztRPsq_uxD4ywazlNwaQoVVg_pQVBgG.RPK9OrDOFRgZgrQU Passcode: xTx2kd!#	***No quiz yet.***
Jan 17 T (Jan 16/M MLK day)	*Repeat slide 19 (Dijkstra) and 20 (Fibonacci)* ***Attendance*** *AI* SEARCH *(From my slides)*: 8-puzzle: https://n-puzzle-solver.appspot.com/ BFS- Uniform Cost; DFS-Depth Limited; Iterative-deepening; Intro to guided search https://fit.zoom.us/rec/share/n_KO-qNpL2owhP4I4XdjLSi8J_NDA0tZkW7-r95-8nSSRhHMXYPpsVFHvZApD4q7.AJTwqlesLZjyjCmP Passcode: hYq0!P7i	*Pair up for the project. (avoid 1 or 3 persons' group).* *Projects will be described below this table.*
Jan 19 R	Heuristic Search A*, (MySlides 15-27) Text Slide Ch4a Start back on MySlides from 27-.., IDA*, etc. InClass Quiz on Search (5 min)	***UG project intermediate submission with due date, below this table*** ***Grad students pair up (or 3 each group?) for the project.*** ***Grad Proposal due date (1-2***

		page/group): ???F on canvas (check for "initial steps" in each project, below) ***You may choose one of these projects (no project is "light") or I may assign one from the choose your own project after talking to me, above. Proposing your own idea is fine too but talk to me first.***
	Updated up to the above date for Spr'23	
Jan 24 T	Local search Work on stack and queue on the recursion tree for Fibonacci computation: show animated steps	***Reminder: If you are registered through the dean-of-students for exam accommodation and need to avail it, let me know by today's class for the upcoming test on the 8th/R***
Jan 26 R	***No lecture, but an on-Canvas test during the class time***	***~10 q/ 20 min regular time***
Jan 31 T	Basics of CNN for grad projects Refresher: My Search slides: 28,32,38,30 Local search: Textslides	**UG project updated below.** **Form UG groups by this week by responding to my Announcement on canvas.** ***Grad project proposal due posted below (more detail later).*** ***Even if you do not have any partner at the moment, select one of the projects, research a bit, and upload proposal on your own*** ***FOR GRADING I USE MY OWN SPREADSHEET. IGNORE AGGREGATES ON CANVAS***
Feb 02 R	*Grad project assignment* *ANN-CNN MyMLslide 31, 33, 39, 42, 45* Local search MySlides (Sl 42--)	***Late Test1 @5pm today*** ***Canvas did not work in class*** *GradProject-1: Proposal* ***Graduate presentations are mandatory for Undergraduate students to attend. I may ask questions on them in any test.***
Feb 07 T	REASONING WITH CONSTRAINTS: Motivating with Map/Graph coloring, Backtracking, Forward Checking (TextSlides) Constraints: MySlides 1-15 *Admissible & Consistent heuristics revision*	*No proposal for UG project. See below for Intermediate report due date.* *Background/Combinatorics quiz answers*

	Discussion on Grad projects	
Feb 09 R	*Partnering issues* *Sudoku Combinatoric (if canvas problem gets solved)* Adversarial search: Textslides (Ch6, up to slide 24)	***UG project Phase 1 (description below this table) submission*** *Little more clarification from a Q/A, below.*
Feb 14 T	*UG Project-1, submit by tonight* *Search Test next week – details later* Automated reasoning / Propositional Logic: Models, (up to text-slide 30).	***My presentation in CS seminar 9/30/F/12pm***
Feb 16 R		
Feb 23 T	Automated reasoning / Propositional Logic: Models, Inferencing, Logical equivalence, Syntax-Semantics-Model, Satisfaction-Entailment-Inference procedure-Validity, Model-checking algorithm, Forward chaining algo, Backward chaining algo. (TextSlides 25-end, Ch-7)	
Feb 25 R	*In-class quiz on Prop Logic* AUTOMATED REASONING: Contd. Unification, Forward chaining algo, Backward chaining algo (sl#40-66) CNF, Resolution Algo (p255), Horn Clause,	**Search-module test, online, in-class, on 10/13/R** ***Ali, please see me after the class.*** *Occur check*: Substitution like x/S(x) not allowed in most logic programming languages, Unification-algo syntactically checks for that and fails in that case. Otherwise, x/S(S…(x)…) infinite looping – source of semi-decidability of FOL Occur check makes an algorithm unsound!
Mar 02, R	**Search-module test, online, in-class, 65min** Question: For the Search Algorithm Test, is the test material on Local search, Informed search, and adversarial search algorithms? Another question, is the test in-person? or can it be taken anywhere, but within class time?	***You may need rough sheets to work on.*** ***Some questions will refer to .pdf files on canvas.*** ***Special students may not have enough extra time available in the***

	My answer: Yes, everything covered under Search. And, also yes on in-person, in-class, and on-line test. You need a device to access Canvas. Question: I was confused about the announcement that was sent out earlier today. Do we need to be in class to take the test today or can we just take it online during class time? My answer: On-line test while sitting in the class.	***class room, as another class will start there. They may take the test outside class room, e.g., at ASC.***
Mar 07 T	**Graduate presentations** *(~10 min each group, have slides/demo, see below)* William & Harold Connor DPLL algorithm Automated Reasoning: First-Order Logic (FOL) (Ch 8) Syntaxes, ForAll and ThereExists quantifiers, writing FOL statements, equivalences (Ch8, up to sl 23)	**Makeup test tonight at 7:30pm, only for students who were excused during the regular test time**
Mar 09 R	**Graduate presentations contd.** *(~10 min each group, have slides/demo, see below)* Automated Reasoning: (Ch 8 up to sl 23, sl 24-bottom: Substitution, sl 26: diagnostic-causal, sl 28: Frame-Quantification-Ramification-etc. problems, ignore rest of the slides.	
Mar 14 T	(Ch 9) inferencing algorithms in FOL: Completeness-Herbrand Universe, Unification, Standardization Apart, Forward and Backward Checking (FC, BC) algorithms for Horn Clause **Graduate presentations contd.**	
Mar 13-17 Spring Break	*--- ---*	
Mar 21 T	*Key concepts*: Unification, Skolemization for quantifiers removal, CNF conversion, … *Additional skill(s)*: Knowledge Engineering or back and forth translations between English	***Let's have a homework, AR-HW-2 on Canvas***

	sentence and FOL sentence. Resolution algorithm for FOL, Examples of logic prog languages **Graduate presentations contd.**	
Mar 23 R	Probabilistic REASONING: Ch13-Textslides-Joint distribution table, conditional probability, simple inferencing: up to sl20 *(Sample exercise Ch13: 13.8, 14, 15)* **Graduate presentations contd.**	**UG Project Phase-2 updated. Due date provided below.**
Mar 28 T	Probabilistic Reasoning (Ch13): deFineti sl10, Start from sl 16 **Graduate presentations contd.**	**>My class page is active now. All slides are there, updated.** ***Next Tuesday, Test-2 on Automated Reasoning*** ***Possible make-up class on Dec 3, Sat***
Mar 30 R	(Ch13) Naïve Bayes and Wumpus-example (Ch14a) Bayesian network, up to sl 9, Sl 22-end **Graduate presentations contd.**	***Test-2 on Automated Reasoning*** ***Same format as before.*** ***After class, take-home, between 6-8pm.*** ***Duration 40min for regular students, <=20 questions*** ***Next class on zoom only due to Hurricane N.*** ***Is there any Physics-minor/major in the class? Do you know one?*** ***ProbReas test on 11/17/R***
Apr 04 T	(Ch14a) Bayesian network, up to sl 9, Sl 22-end **Graduate presentations contd.**	
Apr 06 R	Machine Learning: first, from text-slide Ch18	***Undergraduate project report. Extended, see below***

Apr 11 T	MACHINE LEARNING MySlides: Decision tree, basics	***Test3-Probabilistic Reasoning test (Ch 13, 14a, as covered in class). In-class test, 40 min regular, 14/15 questions, same format as before.***
Apr 13 R	MACHINE LEARNING: linear regression, classification, Perceptron MACHINE LEARNING: ANN from text.	
Apr 18 T	Non-parametric learning kNN variations LSH (start - slide 51), kernel-regression	***FINAL EXAM WILL BE COMPREHENSIVE – all 4 modules, Search, Logic, Probabilistic reasoning, Machine Learning, and Ethics of AI***
Apr 20 R	Contd. ML: Support vector machine Unsupervised learning: Clustering basics: K-means, from https://en.wikipedia.org/wiki/K-means_clustering K-median, Slide 64 hierarchical, density-based, hierarchical	
Apr 25 T (Last day of class 4/26)	***APPARENTLY, NO BACK UP CLASS NEEDED FOR US. I FEEL OBLIGED TO SHOW UP, AND IF THE ROOM IS AVAILABLE, LECTURE ON SOME REVISION. NO ATTENDANCE WILL BE TAKEN. POSSIBLY, ZOOM WILL NOT WORK***	***Graduate Project Final report due on Canvas Dec-6 zipped file (7zip, not .rar): (i) A report between 4-10 pages, including an abstract and references, (ii) Your source code including instructions on how to run and any dependencies and required library details, (iii) if relevant, data files with metadata information (to understand data), and (iv) your updated presentation. All in a zip file, named as: Fa22_one group member's name _project name's signature***
	Lecture until 4pm	**Graduate Project Final Submission due.**

Everyon

	Test4-ML starts at 4pm in-class	
	AI and Ethics (my slides are <u>included</u> in syllabus/tests)	
	Final exam: On-line, comprehensive syllabus.	`//www.fit.edu/registrar/final-examination-schedules/` **Friday, May 5 10:30–12:30 p.m. (p.m.? I suspect they are typos on the web page)**
		<u>Please do the instructor review</u>.

Undergraduate Project: *Pseudo-Sudoku game with variable matrix sizes*
Pseudo-Sudoku (PS) is a simpler version of Sudoku puzzle, but without any 3x3 Block (or Nonet) constraint requirement, as it is in the case of conventional 9x9 Sudoku. However, the board size *nxn* may vary arbitrarily. A valid complete PS board has each row with numbers 1 through *n* (for *n x n* board) without repetition of any number, and so does each of the columns. A given input PS board has some entries (say, *m* such that $m << n^2$) as empty. An algorithm for solving it must fill in these empty items with values satisfying the constraints.

Team up yourselves in pairs. *Suggestion:* Think how to generate PS games for input with chosen (*n,m*) pair of numbers. Better to create valid PS matrices first, and then delete random *m* entries from the matrix.

Assignment: Write *search algorithm(s)* to solve PS games and track your solving-time against the input problem *size* (*n*) and *difficulty* level (*m*) of the problem. [Is there any other way to quantify the difficulty level?] Fancy I/O is not necessary, however, try to make I/O easy enough to understand.

Downloading and using existing code is OK, but not encouraged. Rather, enjoy developing your own AI. If you use code from somewhere, report its source. Understand the code thoroughly, and make at least some modifications in the code, and report that too.

Intermediate submission: At least one algorithm working over n=4. (40% of project grade)
Due: Feb 2, Thursday 11:59pm, on Canvas
Report (one per group): Algorithm/ pseudo-code (not code), editor screen shot showing how your code is running (or compilation screen shot for command line), 3 sample-runs' screen shot, a paragraph of "other" descriptions.

* For the Sudoku project, I have a few questions:
1.) For the I/O input, are we generating a sudoku puzzle and randomly filling it? Or have a text file containing the sudoku puzzle to be read in and solved?
> Have a text file for input, which your program should solve.
2.) What is the difference between compilation screenshots and the three sample-runs screenshots?
> I wanted to see sample runs of input/outputs as your code runs, in your submission file. Not sure how you do it. In "good old days," it would be a screen shot of command mode runs!
3.) Can you elaborate on what the "other description" means? like describing what each function does?
> Anything you want to tell.
4.) For the algorithm, are you expecting a specific kind of algorithm? My partner and I plan to use the backtracking algorithm and wanted to see if we're on the correct path?
Have text file(s) for input and solved file(s) as output.
> Sounds good.

* Are we creating valid sudoku puzzles or are we taking already generated sudoku puzzles and determining if they are valid?
> None. You solve *pseudo*-sudoku puzzle that are partially filled.
I was trying to help and suggest how to create them in the first place, so that you know you are solving one that has a solution.
…is the submission on the ??th a complete model or just an intermediate submission?
> Intermediate! Show that your algorithm solves 3 test cases of size 4x4.

-- *Write group members' names! One submission per group.*

-- *No reference?! Mention if you have written from the scratch.*

-- For *nxn* board, number goes from 1 through *n* (or 0 through n-1), not the fixed 9, as in traditional sudoku puzzle.
Backtracking is a depth first search. Better to write the algorithm in a recursive fashion, unless stack data structure is used.

UG Project Phase-2 Final Submission, Due ??? on Canvas. Implement at least 2 algorithms (including the one you have already implemented in Phase-1) and compare their time-complexities against the problem size, represented by two parameters, *n* and *m* where *n* is the board size and *m* is the number of empty entries. The complexity may be measured with the CPU-time or with the number of steps the algorithm takes. Exclude any input/output or non-algorithmic parts of the code from the complexity measurements.

Plot the time-complexity by varying (*n,m*) over a range, e.g., (n=4, and m=2 out of 4^2), ..,(4, 8), … (n=10, m=2 out of 10^2), …, (n=10, m=50). You choose the range as per your computational capability. *Plot* the complexities for each of your algorithms against *n* and *m.*

You will need to create many input (e.g., 30 for each (*n,m*) pairs, possibly by implementing an input generator. The complexity for each (*n,m*) pair may be averaged over these (e.g., 30) cases. *Note, a random input generator may produce invalid input matrix that should be excluded before you run your algorithms over it*. I am not providing any limit on *n* and *m*, as different computing platforms can handle different ranges. However, use a large enough range to get the feel of *complexity growth* against problem-size (*n,m*).

In Phase-1, I asked for some thought on "difficulty level" of an input for the same (n,m). In other words, do all problems with the same (n,m) run with same complexity? I expect to see that discussion/experiments also.

Submission: a report per group containing (1) two algorithms (pseudo-code), (2) complexity analyses with plots, and (3) discussion. Feel free to extend or refactor your interim report.

Graduate Project:
Submissions include a proposal, an intermediate progress presentation, and the final report. More details will be posted later. I value your efforts and learning, rather than the accuracy of your results. However, good results may lead to satisfaction (and, possible thesis/publication, not to mention better grade) *Do not copy/paraphrase from literature or past projects. When presenting basics, try to provide your own understanding rather than copy-pasted keywords. Do not drag on too much on basics, your own work and results are more important.*

Project proposal due **???/F/2023/6pm.** Team members' names, define the project's scope, describe your current understanding of the project, and especially, what will be your input-output. Provide your data source and/or data generation/augmentation plan (do not underestimate this), expected results and how will you validate accuracy, a few references, possible code-source, etc. Suggested, 1-3 pages, unless figures are added.

Intermediate presentation, ???/T. Per group 10 min. Tutorial on your method (e.g., on ANN/CNN/SOM/etc.); resources: data generation process: demo of running code: expected output and the plan about how you would validate whether the project is successful or not.

IVC SCRIPT
1. Detect the motif on Indus Valley seals using CNN. [Reference: our past work.]
IV seal image -> Motif id.
Learning objective: CNN

2. Given two selected IV fonts, locate them in IV seals. Train with augmented variations of the fonts and validate their recognition in a finite set of IV seal pictures.
Input image -> bounding box
Learning objective: CNN

MEDICAL IMAGING
3. Compare matrix factorization and independent-component analysis over MRI images.

4. Clustering of time series from dynamic images: Ultrasound.

5. Mouse PET images to obesity detection.

6. Organ segmentation using U-net.
Learning objective: U-net

PHYSICS
7. Given a Gaussian distribution, estimate its mean and standard deviation, using ANN.
Vector of real numbers-> two numbers.
Learning objective: Fully connected ANN

BIO-SEQUENCES

7. Density-based clustering a set of bio-sequences from Influenza viruses, with Kobzarenko distance
Learning objective: DBSCAN clustering

8. Density-based clustering a set of bio-sequences from Influenza viruses, with Levenstein distance
Learning objective: DBSCAN clustering

An important notice for Grad students: Program Plan Requirements for New Students or Those Not on File

Graduate policy requires the submission of a program plan by the time you have 9 hours earned or in progress. For most, this will be required sometime in the first semester if registered full time. Future registration may be prevented until an approved plan is on file. With spring registration coming in early November, now is a good time to start working with your advisor to get this prepared and submitted. You can review your holds in PAWS to see if this pertains to you. Below are links to the university templates for MS and Doctoral plans.
https://www.fit.edu/media/site-specific/wwwfitedu/registrar/documents/registrar-forms/MS_prog_plan.pdf
https://www.fit.edu/media/site-specific/wwwfitedu/registrar/documents/registrar-forms/PhD_Prog_plan.pdf

Additionally, if you already have a plan on file and need to make changes, you can use this form (both MS and Doctoral can use this form):
https://www.fit.edu/media/site-specific/wwwfitedu/registrar/documents/registrar-forms/Req_chng_MS_prog.pdf

Lastly, I am available to route your plan if you need assistance with signatures. If you obtain the signatures needed manually, be sure to forward me a copy so that I can have any holds removed accordingly.

Have a great day everyone.

Cheryl Mitravich-Student Coordinator for CES programs
cmitravi@fit.edu
321-674-7047