Intelligent Systems (AI-2)

Computer Science CPSC422, Lecture 1

Jan 11, 2021

People

Instructor

- **Giuseppe Carenini** (carenini@cs.ubc.ca)

Natural Language Processing, Summarization, Discourse Parsing, Preference Elicitation, Explanation, Adaptive

Visualization, Intelligent Interfaces......
Office hour: zoom, TBD
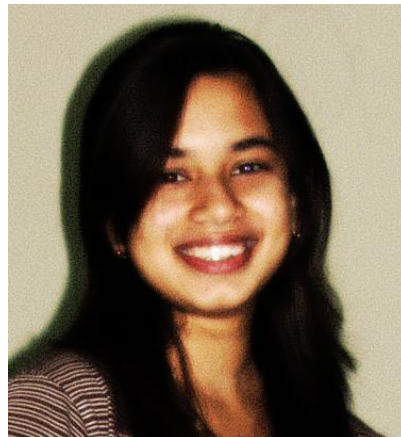
Teaching Assistants (office hours TBD)

Tootooni Mofrad Amirhossein
tootooni@cs.ubc.ca

Deka Namrata
dnamrata@cs.ubc.ca

Ivanova Inna
inna.ivanova@alumni.ubc.ca

Course Essentials(1)

- **Course website:**

www.cs.ubc.ca/~carenini/TEACHING/CPSC422-21/index.html

  - This is where most information about the course will be posted, most handouts (e.g., slides) will be distributed, etc.
  - CHECK IT OFTEN! (draft already available on Canvas)

- **Lectures:**
  - Cover basic notions and concepts known to be hard
  - I will try to post the slides in advance (by 3:30).
  - After class, I will post the same slides inked with the notes I have added in class.
  - Each lecture will end with a set of learning goals:

  Student can….

Course Essentials(2)

Textbook: Selected Chapters from

- **Artificial Intelligence**, 2nd Edition, by Poole, Mackworth. http://people.cs.ubc.ca/~poole/aibook/

Reference (if you want to buy a book in AI this is the one!)

- **Artificial Intelligence: A Modern Approach**, 4th edition, by Russell and Norvig [book webpage on course webpage]

More readings on course webpage…..

Course Essentials(3)

- **Piazza :** discussion board

Signup Link:
piazza.com/ubc.ca/winterterm22020/cpsc422101

- Use the **discussion board** for questions about assignments, material covered in lecture, etc. That way others can learn from your questions and comments!

- Use **email** for private questions (e.g., grade inquiries or health problems).

- **Canvas :** assignments, grades, important links

- **iClicker Cloud** (see instructions on Canvas)   iClicker Cloud

- **AIspace** : online tools for learning Artificial Intelligence http://aispace.org/

  - Under development here at UBC!

  - Already used in cpsc322

Course Elements

- **Practice Exercises:** 0%
- **Assignments:** 15%
- *Research Paper Questions & Summaries 10%*
- **Midterm:** 30%
- **Final:** 45%
- **Review Exam:** 1% bonus
- **Clickers** 3% bonus (1% participation + 2% correct answers)

If your final exam grade is >= 20% higher than your midterm grade:

- Midterm: 15% ⬇
- Final: 60% ⬆

Assignments

- **There will be four assignments in total**
  - They will not necessarily be weighted equally
  - They will be submitted using **Canvas (as single pdf file)**
  - Only the **most recent** submissions will be graded

- **Group work** (same as 322)
  - you **can** work with a partner (and get a small **bonus !**)
    - ✓ Each partnership hands in a **joint** assignment submission with both students' names/IDs
  - you may discuss questions with other students
  - you may not look at or copy each other's written work
  - additional details on the course website

Assignments: Late Days (same as 322)

- **Hand in before class on due day** (on Canvas)
- **You get four late days** ☺
  - to allow you the flexibility to manage unexpected issues
  - additional late days will not be granted except under truly exceptional circumstances
- **A day is defined as:** all or part of a 24-hour block of time beginning at the start of the class on the day an assignment is due
- Applicable to assignments only (**not midterm** or **final)**
- if you've used up all your late days (or don't claim them), **you lose 20% per day**
- **Assignments will not be accepted more than four days late**

Missing Assignments / Midterm / Final

Hopefully late days will cover almost all the reasons you'll be late in submitting assignments.

- However, something more serious like an extended illness may occur ☹

- **For all such cases:** you'll need to provide a note from your doctor, psychiatrist, academic advisor, etc.

- **If you miss:**

  - **an assignment,** your score will be reweighted to exclude that assignment

  - **the midterm,** those grades will be shifted to the final.

  - **the final,** you'll have to write a make-up final as soon as possible.

Paper Readings

- We will read and discuss 3-4 (fairly recent) research papers this term

- You will be required to submit **summaries** of the papers along with **questions** about them

- The goal is to give you practice in reading research papers with a **critical eye**
  - Note possible **weaknesses** in the research or alternative approaches that might be better
  - Connect the **concepts** from the lectures to real-world research
  - What would be the **next step**?

- **Late days cannot be claimed for paper readings!**

How to Get Help?

- Use the course **Piazza discussion board** for questions on course material (so keep reading from it !)
  - *Please check for similar questions before posting*
- If you answer a challenging question you may get **bonus points**! ☺

- Go to **office hours**
  - times already available on course webpage

- Can schedule by appointment if you can document a serious conflict with the official office hours (or if other considerations apply)

Getting Help from Other Students? From the Web? (Plagiarism)

- **It is OK to talk with your classmates about assignments; learning from each other is good**

- **But you must:**

  - Not copy from others (with or without their consent)
  - Write/present your work completely on your own (code questions exception)

- **If you use external source (Web, other students, etc.)** in the assignments. Put material in quotes and cite source!

e.g., *"bla bla bla….."* [wikipedia]

Getting Help from Other Sources? (Plagiarism)

When you are in doubt whether the line is crossed:

- Talk to me or the TA's

- See **UBC official regulations** on what constitutes plagiarism (pointer in course Web-page)

- Ignorance of the rules will not be a sufficient excuse for breaking them

Any unjustified cases will be **severely dealt with** by the **Dean's Office** (that's the official procedure)

- My advice: better to skip an assignment than to have "academic misconduct" recorded on your transcript and additional penalties as serious as expulsion from the university!

To Summarize

- All the course logistics are described in the course Webpage

www.cs.ubc.ca/~carenini/TEACHING/CPSC422-21/index.html

Or WebSearch: Giuseppe Carenini

(And summarized in these slides)

- Make sure you carefully read and understand them!

Let's try
Your UBC-AI Background
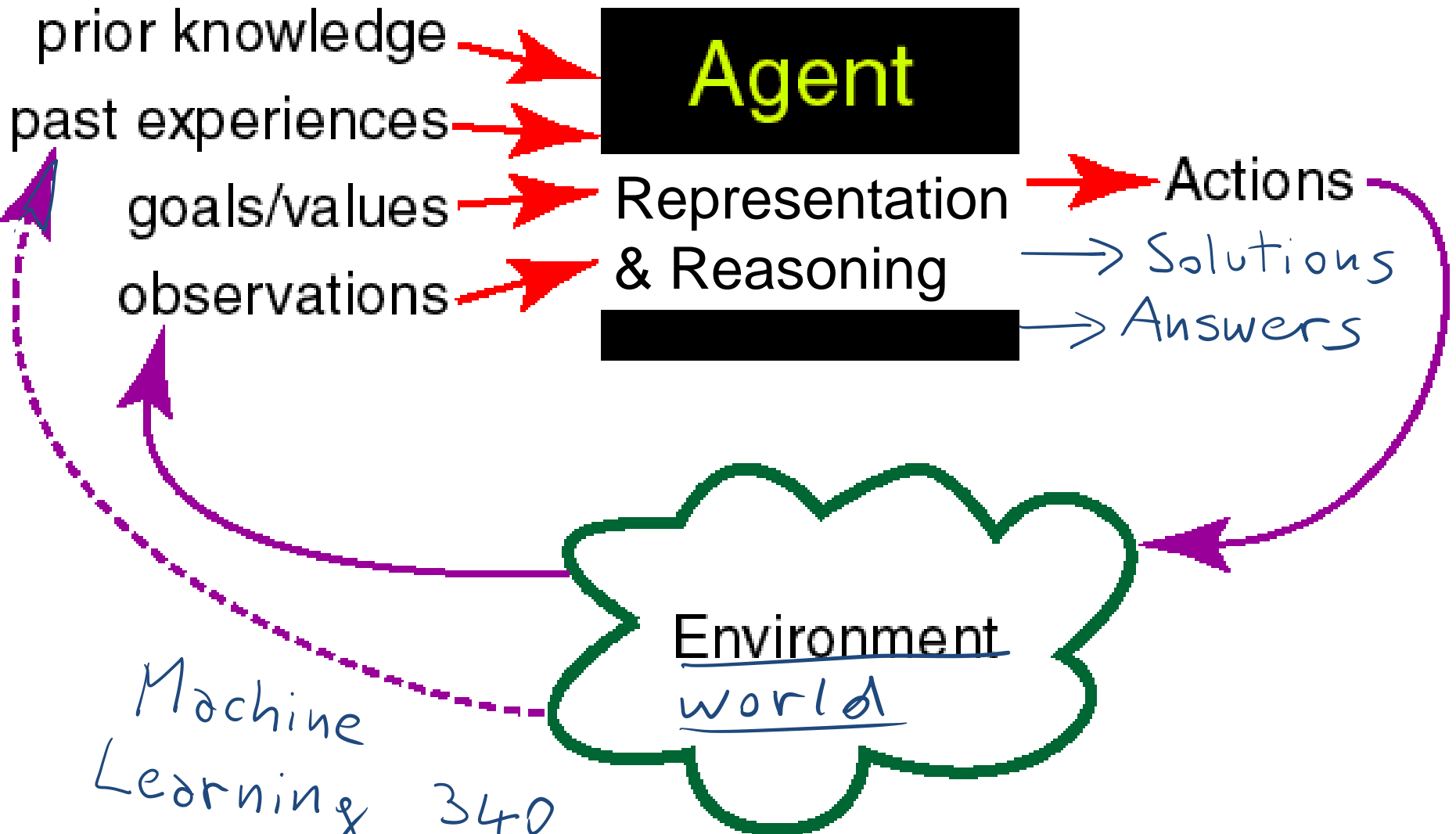
iClicker Cloud

I took 322 within the last 12 months

A. *yes*   **B.** *no*

I took Machine Learning (340) previously

A. *yes*   **B.** *no*

Agents acting in an environment

prior knowledge → Agent

past experiences →

goals/values → Representation & Reasoning → Actions

observations → → Solutions → Answers

Environment world

Machine Learning 340

Cpsc 322 Big Picture

Environment

	Deterministic	Stochastic
Problem		
Constraint Satisfaction	Arc Consistency *Vars +* *Constraints* Search SLS	→for CSP
Query	*Logics* Search →for Inference	*Belief Nets* Var. Elimination *Markov Chains, HMMs*
Planning	*STRIPS* → CSP Search →for complex planning	*Decision Nets* Var. Elimination

Static

Sequential

→ CSP for Inference

Representation

Reasoning
Technique

422 big picture

Deterministic   **Stochastic**

StarAI (statistical relational AI)

Hybrid: Det +Sto
Prob CFG
Prob Relational Models
Markov Logics

Query

Logics
First Order Logics

Ontologies

- **Full Resolution**
- **SAT**

Belief Nets

Approx. : Gibbs

Markov Chains and HMMs

Forward, Viterbi….

Approx. : Particle Filtering

Undirected Graphical Models
Markov Networks
Conditional Random Fields

Planning

Markov Decision Processes and
Partially Observable MDP

- **Value Iteration**
- **Approx. Inference**

Reinforcement Learning
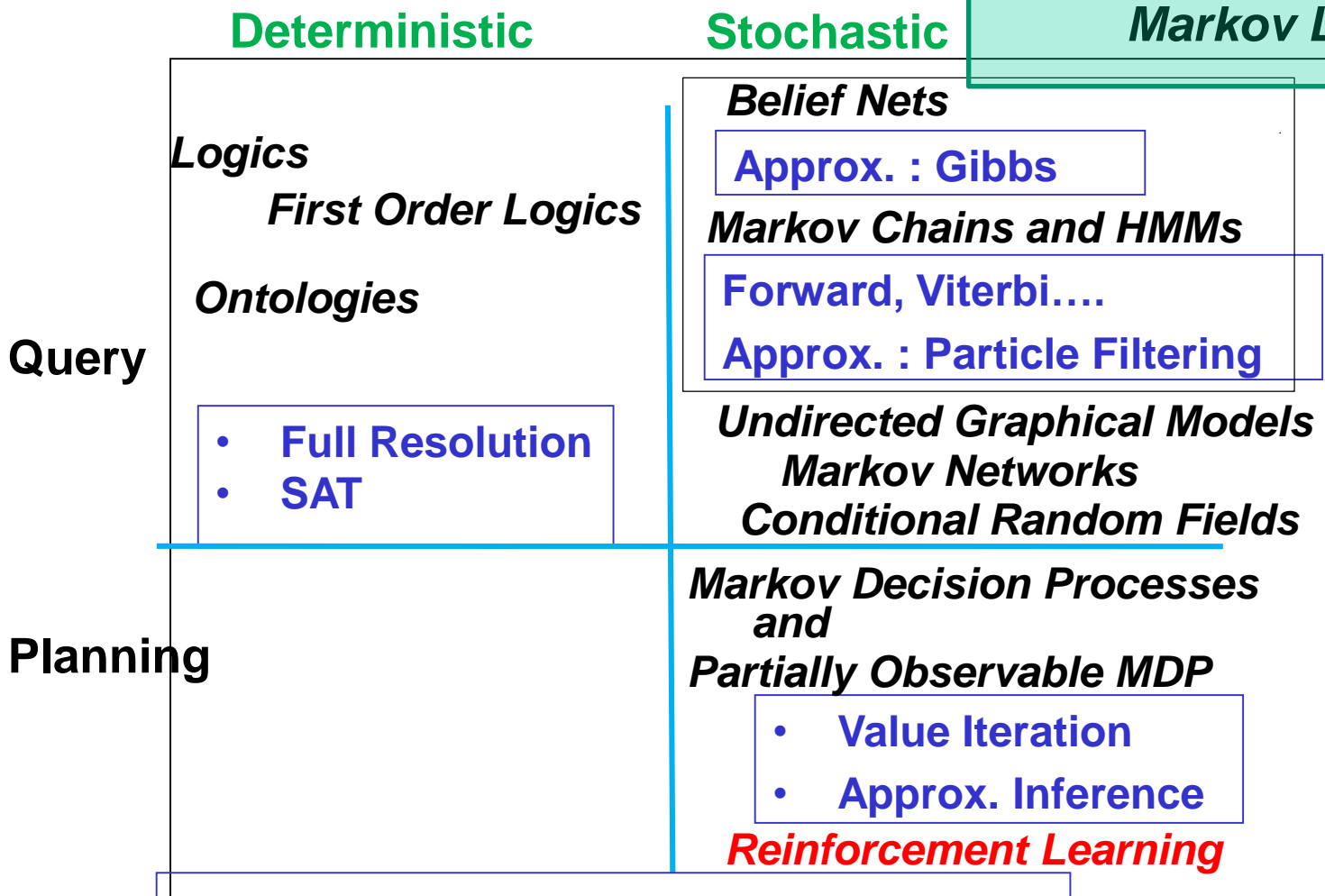
Applications of AI

Representation

Reasoning Technique

Connection with Neural Models / Deep-Learning

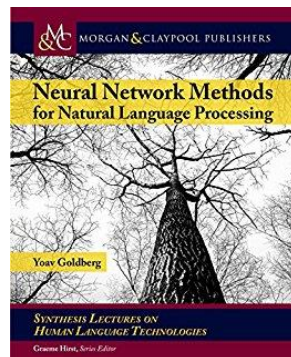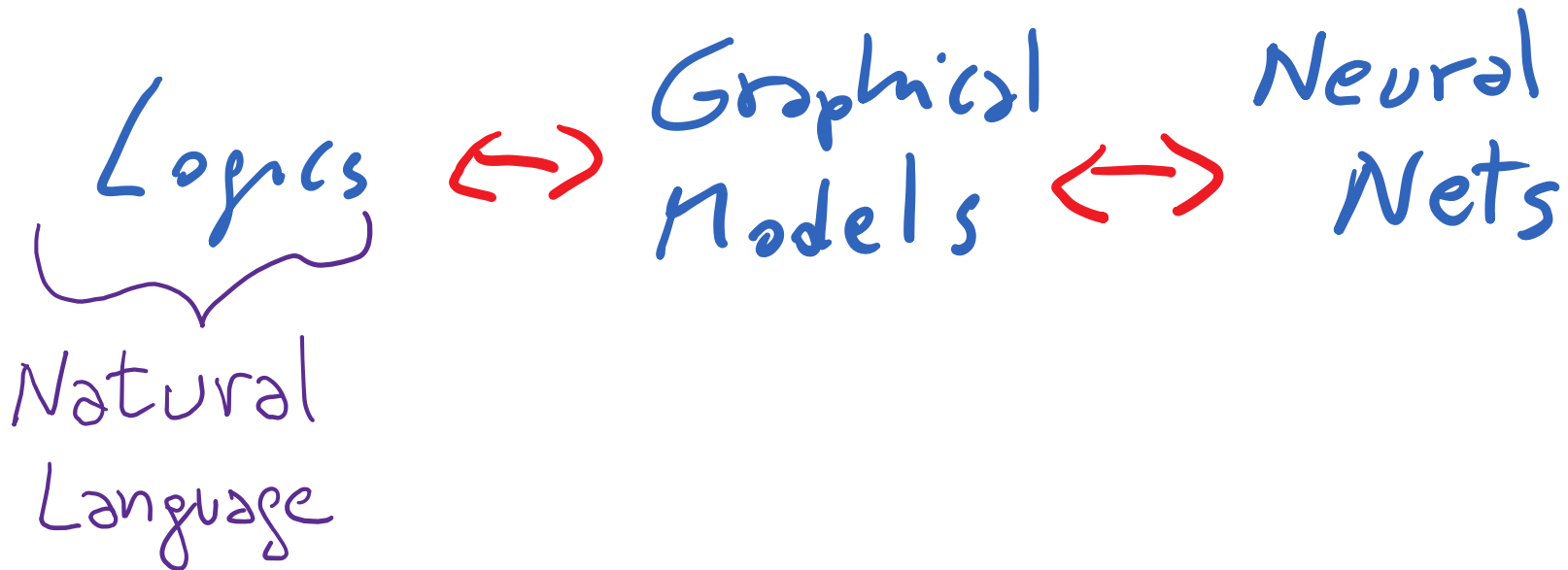NN +Reinforcement Learning: e.g. Alpha-Go (Deep Mind)

http://karpathy.github.io/2016/05/31/rl/
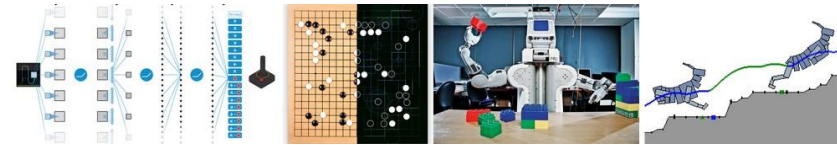
NN (rough prediction) + CRF (refine prediction): Vision (Google paper)
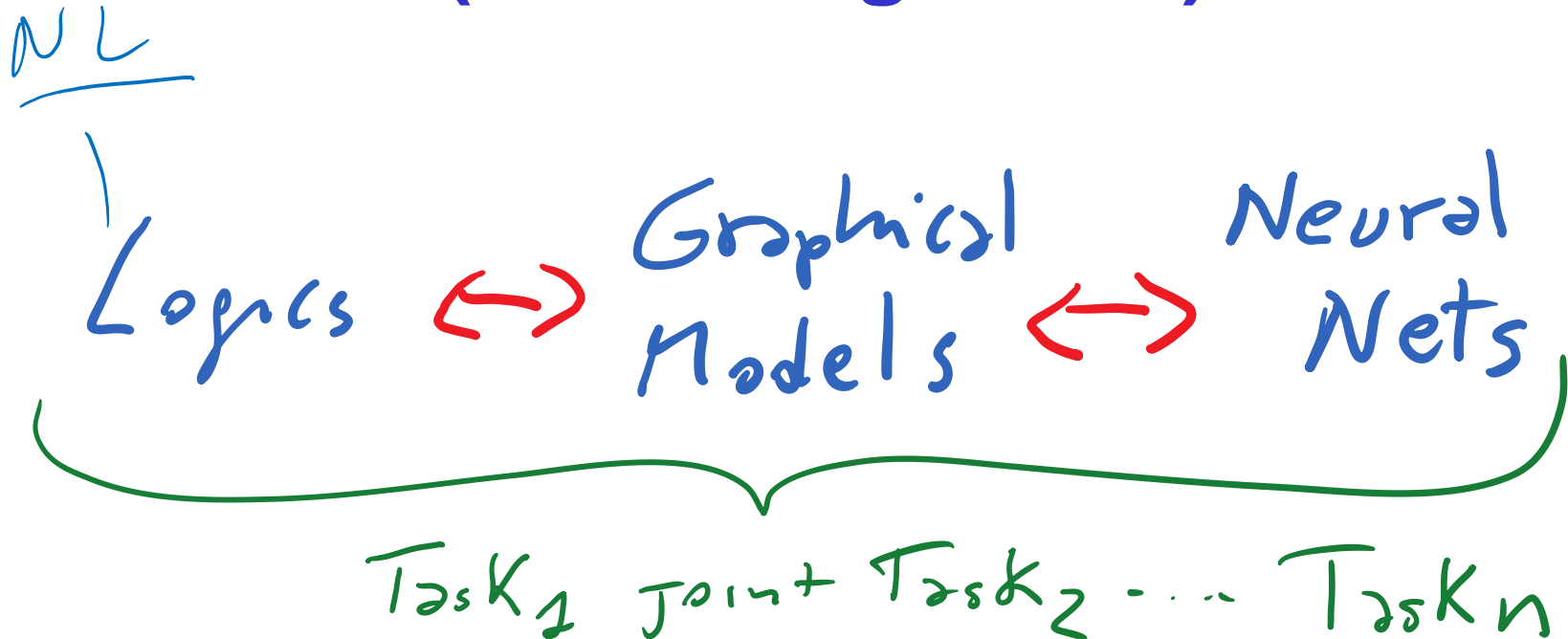
image   CNN output   CRF output
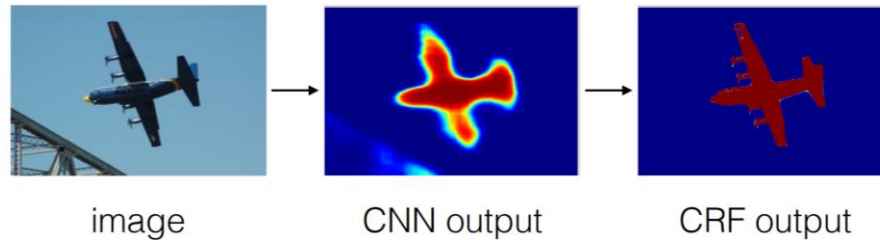
NN + CRF + CKY for NLP (Goldberg book. 2017, Pag 224)

AI Representation & Reasoning & Learning future in a nutshell (according to me)

Logics $\longleftrightarrow$ Graphical Models $\longleftrightarrow$ Neural Nets

Natural Language

AI R&R&L future in a nutshell (according to me)

NL

Logics $\longleftrightarrow$ Graphical Models $\longleftrightarrow$ Neural Nets

$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad}$

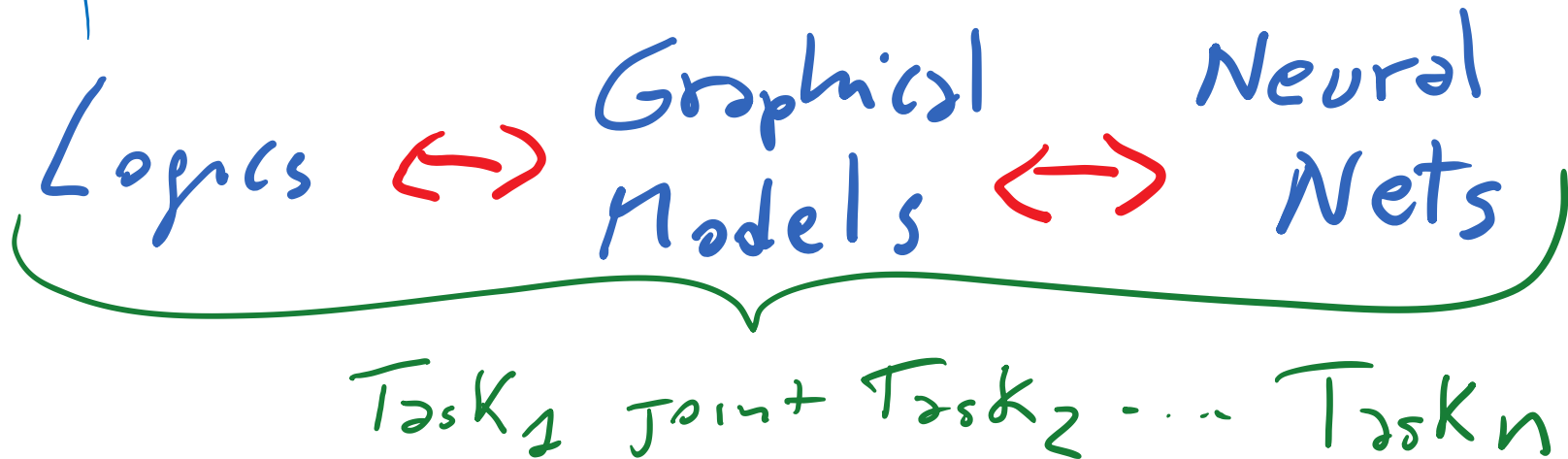Task$_1$ Joint Task$_2$ ... Task$_n$

How to manage multi-tasking for an increasing number of tasks? **Multi-task**

Machine Translation, Text Summarization, Event recognition in videos, Video Captioning

AI R&R&L future in a nutshell (according to me)

NL

Logics $\longleftrightarrow$ Graphical Models $\longleftrightarrow$ Neural Nets

Task$_1$ Joint Task$_2$ ... Task$_n$

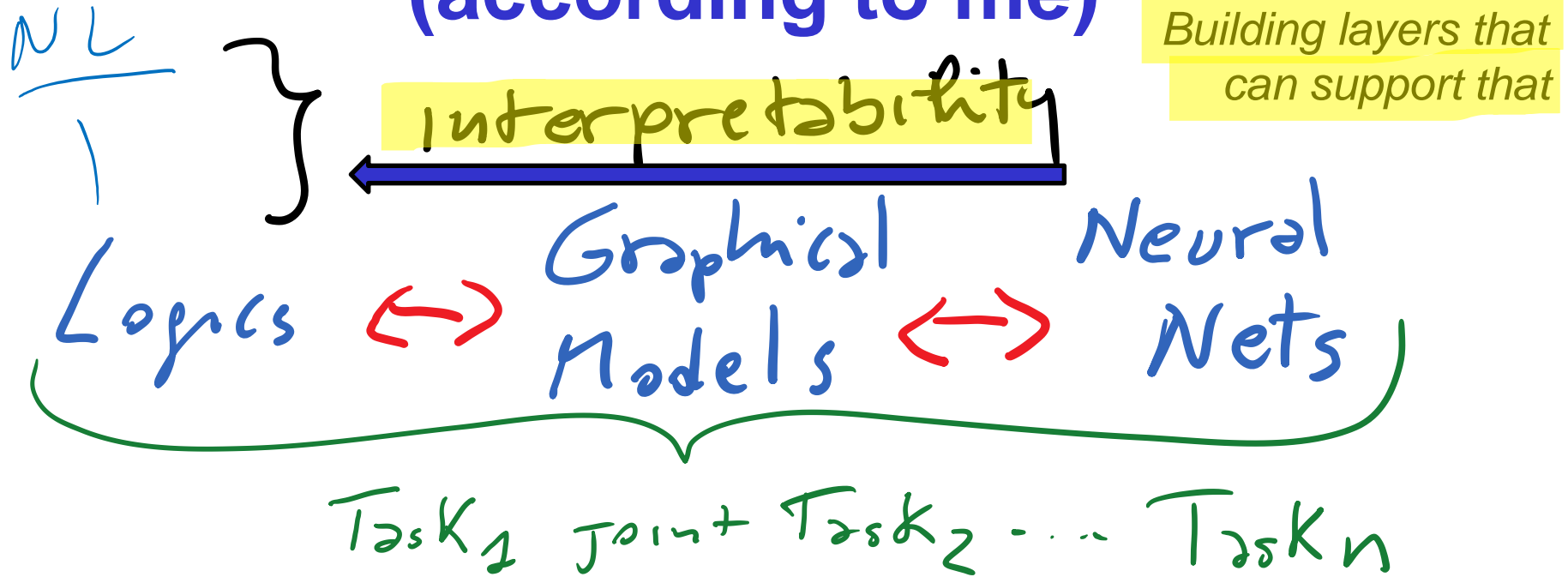How to manage multi-tasking for an increasing number of tasks?

Multi-task R&R&L

In different domains
Medicine, Transportation, Politics, Law…

Machine Translation, Text Summarization, Event recognition in videos

Transfer learning

AI R&R&L future in a nutshell (according to me)

NL

} Interpretability

Building layers that can support that

Logics ↔ Graphical Models ↔ Neural Nets

Task$_1$ Joint Task$_2$ ... Task$_n$

How to manage multi-tasking for an increasing number of tasks?

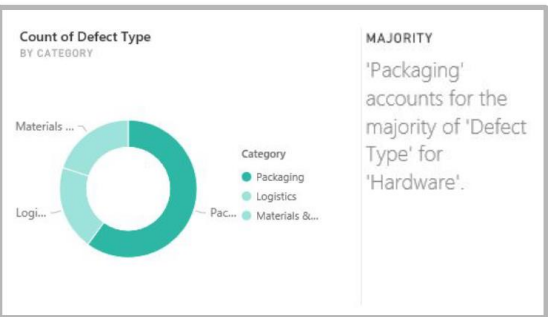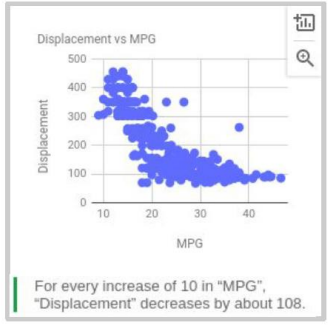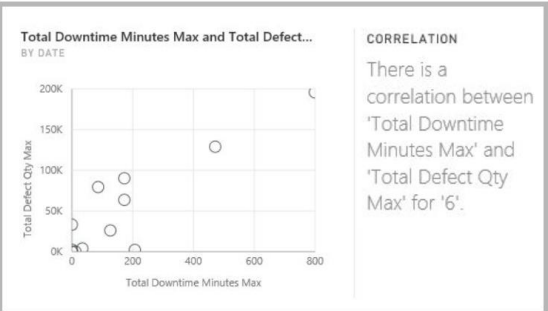**Multi-task**
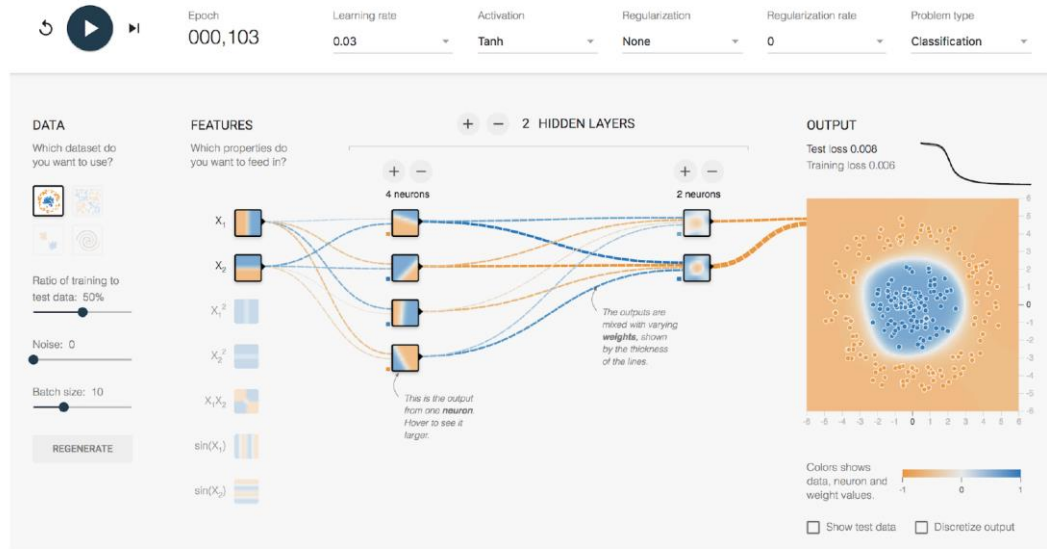
In different domains
Medicine, Transportation, Politics, Law…

Machine Translation, Text Summarization, Event recognition in videos

Transfer learning

Key to interpretability

Language + mathematical abstractions

+ Visualization
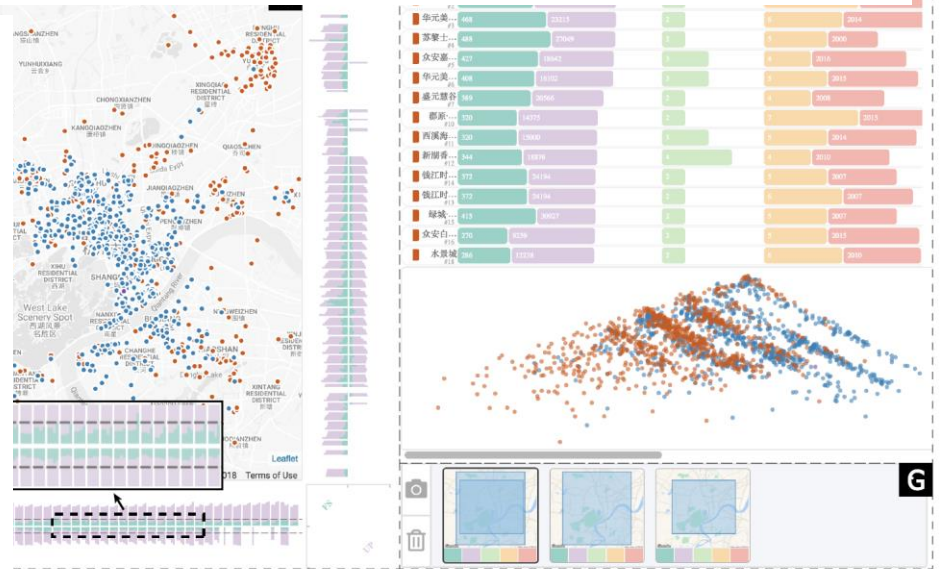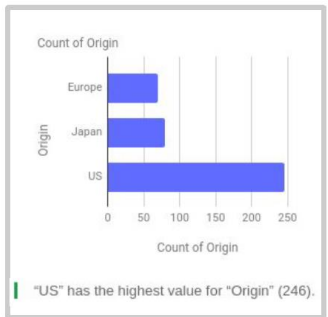
"How is this this course compared to 322?"

- More challenging material (obviously)

- More **mathematical**

  - Lots of probabilistic operations (i.e. sums of products)

  - Important to be able to look at an expression and be able to figure out "what's going on"

- More **programming**

  - Most (if not all) assignments will have a programming component to them

  - Usually you may use the language(s) of your choice

CPSC 322 Review "Exam"

https://forms.gle/ZBVFWxzAZfKCWXFX6

Based on CPSC 322 material

- Logic
- Uncertainty
- Decision Theory

Review material (e.g., 322 slides from 2017):

https://www.cs.ubc.ca/~carenini/TEACHING/CPSC322-17S/index.html

TODO for Wed

- Work on CPSC **322 review exam**

- **Review CPSC 322 material**

- **Read textbook** 9.4, 9.5 (9.5.1 Value of a Policy)

Core Values in 422

- **Conceptual understanding** over **Memorization**
- Sometimes things will be left **underspecified**
    - Eg. I may ask you to solve a problem without telling you the precise steps to follow or what formats to use
    - This is usually **intentional**
- Best answers are:
    - **Correct**
    - **Easy to read**
    - As **short** as possible while including **important** information

From a 2020 PhD thesis…… CRFs

where $s_i=\{w_{i1}, w_{i2}, \ldots, w_{iT_i}\}$, and T_i is the number of tokens of s_i.

To better guide the model with discourse context, we use the shared representations to predict RST relations at the same time. For each paragraph, we run the pre-trained RST parser (Ji and Eisenstein, 2014) and flatten the parse tree to obtain RST relations for each sentence $Y_i=(y_1, \ldots, y_{K_i})$, where K_i is the number of discourse relations in s_i. We then make a label sequence over tokens in the sentence with by placing y at the first word of EDUs and filling up the rest with a *null* relation o: $Y_i' = (o, \ldots, o, \mathbf{y_1}, o, \ldots, \mathbf{y_{K_i}}, o, \ldots, o)$. We incorporate a sequence labeling objective by employing conditional random field (Lafferty et al., 2001) to find the label sequence that maximizes the score function for each sentence s_i: $\mathbf{S}(s_i, Y_i') = \sum_{j=1}^{T_i-1} W_{y_j', y_{j+1}'}^T h_j + b_{y_j', y_{j+1}'}$ where h_j, W and b are the hidden representation of w_{ij}, weight matrix, and the bias vector corresponding to the pair of labels (y_i', y_{i+1}'), respectively. For training, we maximize the conditional likelihood:

$$\mathbb{L}_{CRF} = \mathbf{S}(s_i, y_i') - \sum_{y \in \mathbb{Y}_x} \log \mathbf{S}(s_i, \mathbf{y}) \tag{3.6}$$

where $\mathbb{Y}_x$ represents all possible discourse label sequences. Decoding is done by greedily predicting the output sequence with maximum score. Both training and decoding can be computed using dynamic programming. The final objective is represented as the sum of two objective functions:

$$\mathbb{L}_{\text{disc}} = \mathbb{L}_{s2s} + \alpha * \mathbb{L}_{CRF} \tag{3.7}$$

where α is a scaling parameter to control the impact of CRF objective. The value is chosen empirically by searching based on validation set.

Clickers - Cheating

- **Using another person's clicker**

- **Having someone use your clicker**

is considered **cheating** with the same policies applying as would be the case for turning in illicit written work.