High Performance Computing (HPC) - Oral Exam Prep

Unit 1: Parallel Computing Paradigms

1. What is parallel computing?

Parallel computing is the simultaneous use of multiple processors to solve a computational problem faster.

2. What is Flynn's Taxonomy?

It classifies computer architectures into SISD, SIMD, MISD, and MIMD based on instruction and data streams.

3. What is Amdahl's Law?

Amdahl's Law shows the maximum speedup possible using multiple processors, considering serial portions.

4. What is speedup?

Speedup is the ratio of time taken by a single processor to time taken using multiple processors.

5. What is efficiency?

Efficiency = Speedup / Number of processors; it shows how effectively processors are utilized.

6. Define granularity.

Granularity refers to the amount of computation in relation to communication in a parallel system.

7. What is shared memory architecture?

In shared memory, multiple processors access the same global memory.

8. What is distributed memory?

Each processor has its own private memory and communicates through message passing.

9. What is PRAM?

Parallel Random Access Machine (PRAM) is a theoretical model for designing parallel algorithms.

10. What is SIMD?

Single Instruction, Multiple Data – one instruction operates on multiple data points simultaneously.

1. What is OpenMP?

OpenMP is an API for writing parallel programs using shared memory on multicore systems.

2. What is the #pragma omp parallel directive?

It tells the compiler to execute the code block in parallel using multiple threads.

3. What is the function of omp_get_thread_num()?

It returns the thread ID of the thread executing the code.

4. What is omp_get_num_threads()?

It returns the total number of threads currently in use.

5. What is a critical section in OpenMP?

A critical section ensures that only one thread executes that part of the code at a time.

6. What is the use of parallel for directive?

It divides loop iterations among multiple threads for faster execution.

7. What is dynamic scheduling?

It assigns loop iterations to threads at runtime to balance the workload.

8. What is barrier in OpenMP?

A barrier waits for all threads to reach a certain point before continuing execution.

9. What is thread private variable?

A variable declared private to each thread, avoiding data sharing.

10. Where is OpenMP used?

It is used in scientific simulations, image processing, and large-scale numerical computations.

✓ Unit 3: MPI Programming

1. What is MPI?

Message Passing Interface (MPI) is used for communication between processes in distributed memory systems.

2. What is MPI Init?

It initializes the MPI environment and must be called before any other MPI function.

3. What is MPI_Comm_rank?

It returns the rank (ID) of a process in a communicator group.

4. What is MPI_Comm_size?

It returns the total number of processes involved in the communication.

5. What is point-to-point communication?

Direct communication between two MPI processes using send and receive.

6. What is collective communication?

Communication involving a group of processes, such as broadcast or reduce.

7. What is MPI Barrier?

It synchronizes all processes at a certain point in the code.

8. What is MPI Bcast?

Broadcasts data from one process to all other processes.

9. What is MPI_Sendrecv?

Sends and receives messages in a single function call.

10. Where is MPI used?

Used in high-performance applications like weather forecasting and molecular dynamics.

Unit 4: Parallel Algorithms

1. What is a parallel algorithm?

An algorithm designed to divide tasks among multiple processors for faster execution.

2. What is parallel reduction?

A technique to compute a result (like sum) by reducing values in parallel.

3. What is the parallel version of matrix multiplication?

Matrix rows or cells are distributed among threads/processes to multiply in parallel.

4. What is parallel prefix sum?

It calculates cumulative sums in a parallel way for faster performance.

5. What is a barrier in parallel algorithms?

A synchronization point where all threads wait before proceeding.

6. What is load balancing?

Ensuring all processors perform roughly equal work to avoid delays.

7. What is parallel sorting?

Sorting large datasets using multiple processors working together.

8. Why is parallelism important in search problems?

Parallelism helps search large datasets faster by dividing the task.

9. What is task parallelism?

Different tasks run in parallel, as opposed to the same task on different data.

10. What is the advantage of parallel graph algorithms?

They improve performance for large graphs in social networks or route planning.

✓ Unit 5: CUDA Programming

1. What is CUDA?

CUDA is a parallel computing platform by NVIDIA for programming GPUs.

2. What is a kernel in CUDA?

A kernel is a function executed on the GPU in parallel by multiple threads.

3. What is a thread block?

A group of threads that can cooperate and share memory in CUDA.

4. What is grid in CUDA?

A grid is a collection of thread blocks launched to execute a kernel.

5. What is shared memory in CUDA?

Memory shared among threads in the same block for faster communication.

6. What is global in CUDA?

It defines a function that runs on the GPU and is called from the CPU.

7. What is memory coalescing?

Optimizing memory access so that threads access consecutive memory locations.

8. What is warp?

A group of 32 threads executed together in CUDA.

9. Why use GPU programming?

GPUs offer massive parallelism, speeding up data-heavy tasks like deep learning and simulations.

10. Give one application of CUDA.

Used in video rendering, image processing, and scientific computing.

Deep Learning – Oral Exam Prep (Unit-Wise)

Unit 1: Foundations of Deep Learning

1. What is deep learning?

Deep learning is a subset of machine learning that uses neural networks with multiple layers to learn patterns from data.

2. What is supervised learning?

Supervised learning uses labeled data to train models, where input-output pairs are already known.

3. What is the bias-variance tradeoff?

High bias leads to underfitting and high variance leads to overfitting; the tradeoff helps balance model accuracy.

4. What are hyperparameters?

Hyperparameters are external configurations like learning rate, batch size, and epochs that control training.

5. What causes overfitting?

Overfitting happens when the model learns noise in training data instead of general patterns.

6. What is regularization?

Regularization methods like L1 and L2 reduce model complexity to avoid overfitting.

7. Give one limitation of machine learning.

Traditional ML struggles with raw data and often requires feature engineering.

8. What is the role of TensorFlow and PyTorch?

They are popular libraries used for building and training deep learning models.

9. Give one real-life application of deep learning.

Facial recognition and voice assistants like Siri use deep learning.

10. Who developed AlphaGo?

DeepMind developed AlphaGo, which used deep reinforcement learning.

1. What is a perceptron?

A perceptron is a basic unit of a neural network that computes weighted inputs plus bias and applies activation.

2. What is the role of activation functions?

Activation functions introduce non-linearity, allowing the model to learn complex patterns.

3. Name any two activation functions.

ReLU and Sigmoid are commonly used activation functions.

4. What is backpropagation?

Backpropagation adjusts weights using gradients from the loss function to minimize errors.

5. What is gradient descent?

It's an optimization algorithm used to minimize loss by updating weights step-bystep.

6. What is the vanishing gradient problem?

It occurs when gradients become too small during backpropagation, making learning slow.

7. What are loss functions used for?

They calculate the difference between predicted and actual outputs to guide learning.

8. What is the purpose of hidden layers?

They allow the network to extract higher-level features from input data.

9. What is PyTorch?

PyTorch is a flexible deep learning framework preferred for research and experimentation.

10. What is sentiment analysis?

It's the use of deep learning to classify emotions or opinions in text data.

Unit 3: Convolutional Neural Networks (CNN)

1. What is a CNN?

A CNN is a neural network designed to process image data using filters.

2. What is padding in CNN?

Padding adds extra pixels to input to preserve output size after convolution.

3. What is the ReLU function?

ReLU stands for Rectified Linear Unit; it outputs the input if positive, else zero.

4. What is pooling in CNN?

Pooling reduces the spatial dimensions of the image, preserving features and reducing computation.

5. What is the function of convolutional layers?

They extract features from input data using filters.

6. What is the use of fully connected layers?

They connect every neuron to every other in the previous layer for classification.

7. What is stride?

Stride is the number of pixels the filter moves during convolution.

8. What is Local Response Normalization?

It normalizes neuron outputs, improving generalization.

9. Which network made CNN popular?

AlexNet, a deep CNN, won the 2012 ImageNet competition.

10. What is VGGNet?

VGGNet is a CNN model using very small (3×3) filters with deep architecture.

Unit 4: Recurrent Neural Networks (RNN)

1. What is an RNN?

RNNs are neural networks for sequential data like time series or text.

2. What is the difference between CNN and RNN?

CNNs process spatial data, RNNs handle sequential/time-based data.

3. What is backpropagation through time (BPTT)?

BPTT is an extension of backpropagation used in training RNNs.

4. What is an LSTM?

LSTM (Long Short-Term Memory) solves long-term dependency problems in RNNs.

5. What is a bidirectional RNN?

It processes data in both forward and backward directions.

6. What is an encoder-decoder architecture?

It maps input sequences to output sequences, used in translation.

7. What are vanishing gradients in RNNs?

It's when gradients become too small, preventing learning of long-term patterns.

8. What is Echo State Network?

It's a special RNN with a large fixed recurrent layer and trainable output.

9. What is recursive neural network?

It processes hierarchical input structures like parse trees.

10. Where are RNNs used?

Applications include speech recognition, language modeling, and music generation.

✓ Unit 5: Deep Generative Models

1. What is a GAN?

A GAN is a deep learning model where a generator creates data and a discriminator judges it.

2. What is the goal of the generator?

To create realistic data that can fool the discriminator.

3. What is the goal of the discriminator?

To differentiate between real and fake data.

4. Name a use of GANs.

GANs are used for fake image generation and style transfer.

5. What is a Boltzmann Machine?

It is a stochastic, generative neural network that learns probability distributions.

6. What are Deep Belief Networks?

They are stacked restricted Boltzmann machines used for unsupervised learning.

7. What are the types of GANs?

Types include DCGAN, CycleGAN, Conditional GAN, and StyleGAN.

8. What is adversarial training?

It's the process of training GANs where generator and discriminator compete.

9. What is mode collapse in GANs?

When the generator produces limited types of outputs, ignoring variety.

10. How can GANs detect fake images?

The discriminator is trained to classify fake and real images, making it suitable for forgery detection.

Unit 6: Reinforcement Learning

1. What is reinforcement learning?

RL is learning by interacting with an environment to maximize cumulative rewards.

2. What is MDP?

Markov Decision Process models RL using states, actions, rewards, and transitions.

3. What is Q-learning?

It's a model-free RL algorithm that learns action values without knowing the model.

4. What is a policy in RL?

It is the strategy the agent uses to decide actions based on states.

5. What is the role of reward?

Reward provides feedback on how good an action is in a given state.

6. What is exploration vs exploitation?

Exploration tries new actions, while exploitation uses the best known actions.

7. What are Deep Q Networks (DQN)?

DQN combines Q-learning with deep neural networks for complex environments.

8. What are some applications of RL?

Used in gaming (AlphaGo), robotics, and autonomous systems.

9. What is the Tic-Tac-Toe RL example?

It shows how RL can learn strategies by playing and improving from feedback.

10. What is a challenge in RL?

Sparse rewards and high-dimensional state spaces make learning difficult.