WHAT IS CLAIMED IS:

1. A method performed by one or more computers, the method comprising:

obtaining netlist data for a computer chip, wherein the netlist data specifies a connectivity on a computer chip between a plurality of nodes that each correspond to one or more of a plurality of integrated circuit components of the computer chip; and

generating a computer chip floorplan that places each node in the netlist data at a respective position on the surface of the computer chip, comprising placing a respective node at each time step in a sequence comprising a plurality of time steps, the placing comprising, for each particular time step of the plurality of time steps:

generating an input representation for the particular time step that comprises (i) partially placed chip data characterizing a partially placed chip that indicates the placement of the nodes that have already been placed at time steps preceding the particular time step in the sequence and (ii) current node data characterizing the node to be placed at the particular time step;

processing the input representation using a node placement neural network having a plurality of network parameters, wherein the node placement neural network is configured to process the input representation in accordance with current values of the network parameters to generate a score distribution over a plurality of positions on the surface of the computer chip; and

assigning the node to be placed at the particular time step to a position from the plurality of positions using the score distribution.

2. The method of claim 1, wherein assigning the node to a position from the plurality of positions using the score distribution comprises:

generating a modified score distribution that sets the score for each position at which a node has already been placed to zero; and

assigning the node using the modified score distribution.

3. The method of claim 2, wherein assigning the node using the modified score distribution comprises:

assigning the node to the position having the highest score in the modified score distribution.

4. The method of claim 2, wherein assigning the node using the modified score distribution comprises:

sampling a position from the modified score distribution, and assigning the node to the sampled position.

5. The method of any preceding claim, wherein the plurality of positions comprise grid squares from an N x M grid overlaid over the surface of the computer chip.

6. The method of any preceding claim, wherein the current node data characterizing the node to be placed at the particular time step comprises data characterizing the connectivity of the node to be placed at the particular time step specified by the netlist data.

7. The method of any preceding claim, wherein the input representation comprises netlist features that characterize the netlist data.

8. The method of claim 7, wherein the netlist features comprise a respective feature vector for each node in the netlist.

9. The method of claim 8, wherein the feature vector for each node identifies a size of the node.

10. The method of claim 8 or claim 9, wherein, for each node that has already been placed at an earlier time step in the sequence, the feature vector for the node identifies the position at which the node was placed.

11. The method of any preceding claim, wherein the current node data characterizing the node to be placed at the particular time step comprises an embedding of the node to be placed at the particular time step.

12. The method of any preceding claim, wherein the partially placed chip data comprises an image of the surface of the partially placed chip that depicts the positions at which the nodes at time steps preceding the particular time step in the sequence have been

placed.

13. The method of claim 12, wherein, in the image, each position of the plurality of positions is associated with a density value that reflects an occupancy of the position in the partially placed chip.

14. The method of any preceding claim, wherein the node placement neural network comprises:

an input subnetwork configured to, at each particular time step, process the input representation to generate an embedding of the input representation; and

a policy subnetwork configured to, at each particular time step, process the embedding to generate the score distribution.

15. The method of claim 14, when claim 14 is dependent on any one of claims 12 or 13, wherein the input subnetwork comprises a convolutional neural network configured to generate an embedding of the partially placed chip data.

16. The method of claim 14, when claim 14 is dependent on any one of claims 7-10, wherein the input subnetwork comprises a graph convolutional neural network configured to generate an embedding of the current node data and the netlist feature data.

17. The method of any one of claims 14-16, wherein the node placement neural network comprises:

a value subnetwork configured to, at each particular time step, process the embedding to generate a value estimate that estimates a value of a current state of the floorplan as of the particular time step.

18. The method of any preceding claim, further comprising:

determining the current values of the network parameters by training the node placement neural network through reinforcement learning to generate probability distributions that maximize a reward function that measures a quality of floorplans generated using the node placement neural network.

19. The method of claim 18, wherein the reward function includes a wire length measure that is higher when wire length between nodes on the surface of the chip is shorter.

20. The method of any one of claims 17 or 18, wherein the reward function includes a congestion measure that is higher when congestion on the surface of the computer chip is lower.

21. The method of claim 20, wherein the congestion measures track a density of wires across horizontal and vertical edges of the surface.

22. The method of any one of claims 18-21, wherein the training of the node placement neural network is distributed across a plurality of synchronous workers.

23. The method of any one of claims 18-22, wherein the training of the node placement neural network is distributed across a plurality of asynchronous workers that each add trajectories to an experience replay memory.

24. The method of any one of claims 18-23, further comprising:

generating a final computer chip floorplan by performing simulated annealing on the reward function starting from the computer chip floorplan generated using the node placement neural network.

25. The method of claim 24, wherein performing simulated annealing comprises applying a hill climbing algorithm.

26. One or more computer-readable storage media storing instructions that when executed by one or more computers cause the one or more computers to perform the respective operations of any one of the methods of any of the preceding claims.

27. A system comprising one or more computers and one or more storage devices storing instructions that when executed by one or more computers cause the one or more computers to perform the respective operations of any one of the methods of any of claims 1 to 25.

28. A method of producing a computer chip, comprising generating a computer chip floorplan by a method according to any of claims 1 to 25, and fabricating a computer chip according to the generated computer chip floorplan.