What Is Claimed Is:

1. A speech processing system for communicating with a user, comprising:

an input interface configured to receive a speech signal;

a first unified neural network comprising an automatic speech recognition (ASR) section and a natural language understanding (NLU) section, the first unified neural network configured to process the speech signal to extract one or more of intents and entities, wherein a functionality of the ASR section and a functionality of the NLU section are both accomplished by the first unified neural network;

a dialogue manager configured to analyze the one or more of intents and entities to generate a dialogue response; and

a second unified neural network comprising a natural language generator (NLG) section and a text-to-speech (TTS) section, the second unified neural network configured to generate a speech output corresponding to the dialogue response for the user, wherein a functionality of the NLG section and a functionality of the TTS section are both accomplished by the second unified neural network.

2. The speech processing system according to claim 1, wherein the ASR section of the first unified neural network is configured to convert the speech signal into a first network state and the NLU section of the first unified neural network is configured to extract the one or more of intents and entities from the first network state.

3. The speech processing system according to claim 2, wherein the NLG section of the second unified neural network is configured to generate a second network

state corresponding to the dialogue response and the TTS section of the second unified neural network is configured to convert the second network state into the speech output.

4. The speech processing system according to claim 1, wherein the input interface is further configured to receive one or more events and transmit the one or more events to the dialogue manager.

5. The speech processing system according to claim 4, wherein the dialogue manager is further configured to generate one or more control parameters for the first unified neural network based on the one or more events, wherein the first unified neural network is configured to implement one or more models based on the one or more control parameters.

6. The speech processing system according to claim 1, wherein the dialogue manager is further configured to fetch data from an external database for analyzing the one or more of intents and entities to generate the dialogue response.

7. The speech processing system according to claim 1, wherein the first unified neural network and the second unified neural network are configured to implement at least one of one or more bi-directional Long Short Term Memory (LSTM) neural networks and one or more transformer neural networks.

8. The speech processing system according to claim 1, wherein the first unified neural network is further configured to extract a semantic relationship between the one or more of intents and entities; and the second unified neural network

is further configured to analyze the semantic relationship between the one or more of intents and entities to generate the speech output corresponding to the dialogue response for the user.

9. The speech processing system according to claim 1, wherein the entities are composite entities.

10. A speech processing system for communicating with a user, comprising:

an input interface configured to receive a speech signal; and

a single unified neural network, wherein the single unified neural network includes a plurality of sections each with a respective different functionality accomplished by the single unified neural network, the single unified neural network configured to:

process the speech signal to extract one or more of intents and entities;

analyze the one or more of intents and entities to generate a dialogue response; and

generate a speech output corresponding to the dialogue response for the user.

11. The speech processing system according to claim 10, wherein the single unified neural network comprises:

an automatic speech recognition (ASR) section of the plurality of sections configured to convert the speech signal into a first network state;

a natural language understanding (NLU) section of the plurality of sections configured to extract the one or more of intents and entities from the first network state;

a natural language generator (NLG) section of the plurality of sections configured to generate a second network state corresponding to the dialogue response; and

a text-to-speech (TTS) section of the plurality of sections configured to convert the second network state into the speech output.

12. The speech processing system according to claim 10, wherein the single unified neural network comprises a dialogue manager section of the plurality of sections configured to fetch data from an external database for analyzing the one or more of intents and entities to generate the dialogue response.

13. The speech processing system according to claim 10, wherein the single neural network is further configured to process the speech signal to extract a semantic relationship between the one or more of intents and entities, wherein the single neural network is further configured to analyze the semantic relationship between the one or more of intents and entities to generate the dialogue response.

14. The speech processing system according to claim 10, wherein the entities are composite entities.

15. A computer implemented method for speech processing, comprising:

receiving a speech signal;

processing, by a first unified neural network comprising an automatic speech recognition (ASR) section and a natural language understanding (NLU) section, the speech signal to extract one or more of intents and entities, wherein

a functionality of the ASR section and a functionality of the NLU section are both accomplished by the first unified neural network;

analyzing the one or more of intents and entities to generate a dialogue response; and

generating, by a second unified neural network comprising a natural language generator (NLG) section and a text-to-speech (TTS) section, a speech output corresponding to the dialogue response for the user, wherein a functionality of the NLG section and a functionality of the TTS section are both accomplished by the second unified neural network.

16. The computer implemented method according to claim 15, further comprising converting, by the ASR section of the first unified neural network, the speech signal into a first network state and extracting, by the NLU section of the first unified neural network, the one or more of intents and entities from the first network state.

17. The computer implemented method according to claim 16, further comprising generating, by the NLG section of the second unified neural network, a second network state corresponding to the dialogue response and converting, by the TTS section of the second unified neural network, the second network state into the speech output.

18. The computer implemented method according to claim 15, wherein the entities are composite entities.

19. A computer implemented method for speech processing, comprising:

receiving a speech signal;

processing, by a single unified neural network comprising an automatic speech recognition (ASR) section and a natural language understanding (NLU) section, the received speech signal to extract one or more of intents and entities, wherein a functionality of the ASR section and a functionality of the NLU section are both accomplished by the single unified neural network;

analyzing, by the single unified neural network, the one or more of intents and entities to generate a dialogue response; and

generating, by the single unified neural network comprising a natural language generator (NLG) section and a text-to-speech (TTS) section, a speech output corresponding to the dialogue response for the user, wherein a functionality of the NLG section and a functionality of the TTS section are both accomplished by the single unified neural network.

20. The computer implemented method according to claim 19, further comprising:

converting, by the ASR section of the single unified neural network, the speech signal into a first network state;

extracting, by the NLU section of the single unified neural network, the one or more of intents and entities from the first network state;

generating, by the NLG section of the single unified neural network, a second network state corresponding to the dialogue response; and

converting, by the TTS section of the single unified neural network, the second network state into the speech output.

21. The computer implemented method according to claim 19, further comprising fetching, by a dialogue manager section of the single unified neural network, data from an external database for analyzing the one or more of intents and entities to generate the dialogue response.

22. The computer implemented method according to claim 19, wherein the entities are composite entities.

23. A speech processing system for communicating with a user, including one or more processors and one or more memories configured to perform operations comprising:

receiving a speech signal;

processing, by a unified neural network comprising an automatic speech recognition (ASR) section and a natural language understanding (NLU) section, the speech signal to extract one or more of intents and entities, wherein a functionality of the ASR section and a functionality of the NLU section are both accomplished by the unified neural network; and

analyzing the one or more of intents and entities to generate a dialogue response.

24. The speech processing system according to claim 23, wherein the operations further comprise converting the speech signal into a network state and extracting the one or more of intents and entities from the network state.

25. The speech processing system according to claim 23, wherein the operations further comprise generating a natural language text response corresponding to the dialogue response and converting the natural language text response into the speech output.

26. The speech processing system according to claim 23, wherein the entities are composite entities.

27. A speech processing system for communicating with a user, comprising:

an input interface configured to receive a speech signal;

an automatic speech recognition (ASR) based neural network configured to convert the speech signal into text information;

a natural language understanding (NLU) based neural network configured to extract one or more of intents and entities from the text information;

a dialogue manager configured to analyze the intents and entities to generate a dialogue response; and

a unified neural network, the unified neural network comprising a natural language generator (NLG) section and a text-to-speech (TTS) section configured to generate a speech output corresponding to the dialogue response for the user, wherein a functionality of the NLG section and a functionality of the TTS section are both accomplished by the unified neural network.

28. The speech processing system according to claim 27, wherein the NLG section is further configured to generate a network state corresponding to the dialogue response and wherein the TTS section is further configured to convert the network state into the speech output.

29. The speech processing system according to claim 27, wherein the entities are composite entities.