

Dialogue Modeling for Virtual Humans

Sudeep Gandhe | 09/24/2008



The projects or efforts depicted were or are sponsored by the U.S. Army Research, Development, and Engineering Command (RDECOM), and/or the US Army Research Institute. The content or information presented does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.



What is Dialogue?

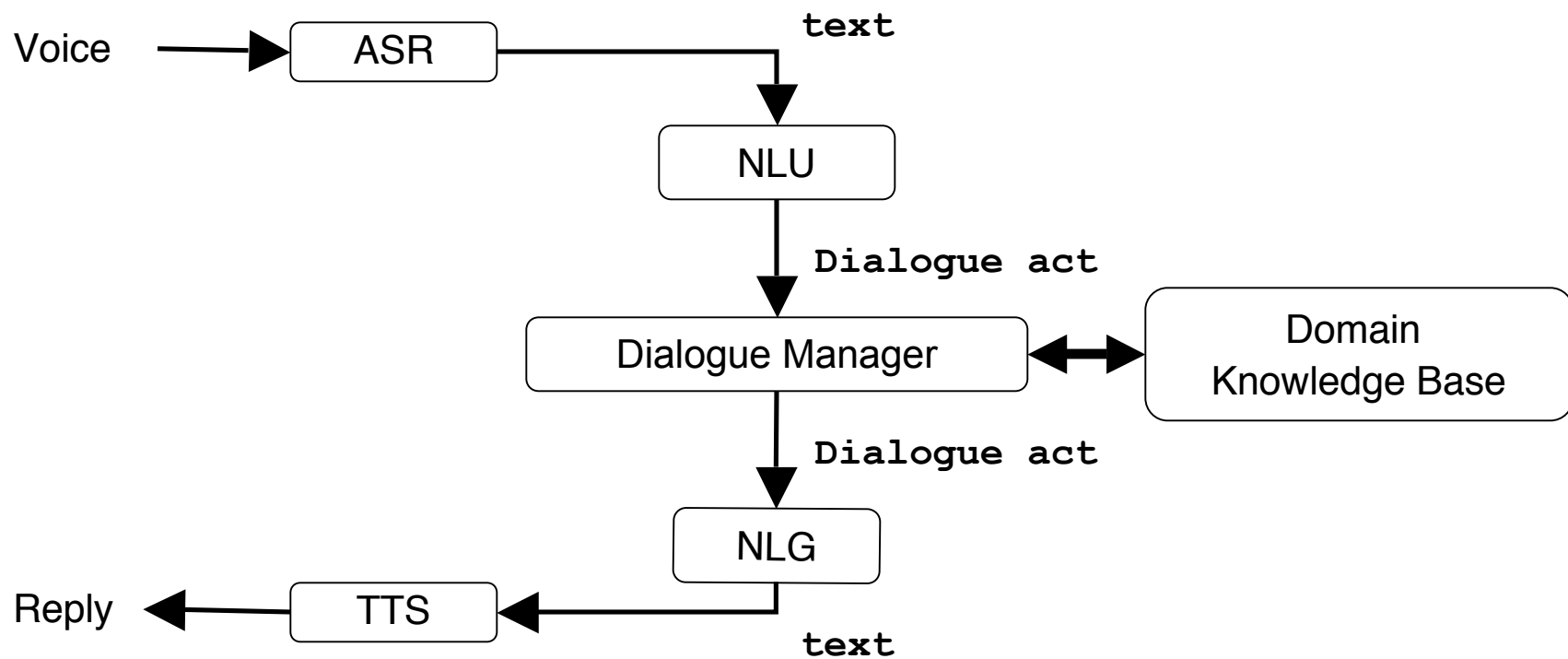
- **Communication**
 - Multiple participants
 - Multiple contributions
 - Coherent
 - Interactive
- **Modalities**
 - Input: Speech, text, sketch, menu, gestures
 - Output: Speech, text, graphical displays, embodied character
- **Dialogue systems**
 - Information exchange
 - Instruction giver
 - Tutor
 - Collaborative partner
 - Conversation partner



What is Dialogue Modeling/Management?

- **Progressively tracking the state of the dialogue**
 - Update the dialogue state with each utterance
 - Optionally track multiple hypotheses to handle uncertainty
- **Providing a context for interpretation of the input utterance(s)**
 - Ellipsis
 - Anaphora
- **Selecting the content and type of the output utterance(s)**
 - May need to connect to external database, Ontology/ domain knowledge

Dialogue System Architecture



What is Speech Act?

- **Speaking is Acting (Austin, 1962)**
- **Speech Act**
 - Locutionary act (act of saying something)
 - Act of producing certain noise, choosing certain words
 - **Illocutionary act** (act in saying something)
 - Request, propose, accept, refuse, etc.
 - Perlocutionary act (act by saying something)
 - Persuasion, surprise, etc.
- **We are interested in illocutionary acts**
 - Illocutionary force + propositional content
- **E.g. from SASO framebank**

We will have to move the hospital .	
S.mood	declarative
S.sem.task	move-clinic
S.sem.speechact.type	statement
S.sem.type	event
S.sem.modal.deontic	must
S.sem.agent	we
S.sem.event	move
S.sem.theme	hospital
S.sem.time	future

What is Dialogue act?

Turn-taking	take-turn, keep-turn, release-turn, assign-turn
Grounding	Initiate, continue, ack, repair, request-repair, request-ack, etc.
Core Speech acts	inform, question, request, accept, reject, etc.
Argumentation	elaborate, summarize, convince, etc.

- **How to get the dialogue act from surface text? (NLU)**
- **Keyword spotting, grammars, statistical methods (classifier)**
- **Need training data**
 - <surface text, dialogue act> examples

Adjacency Pairs (Schegloff & Sacks, 1973)

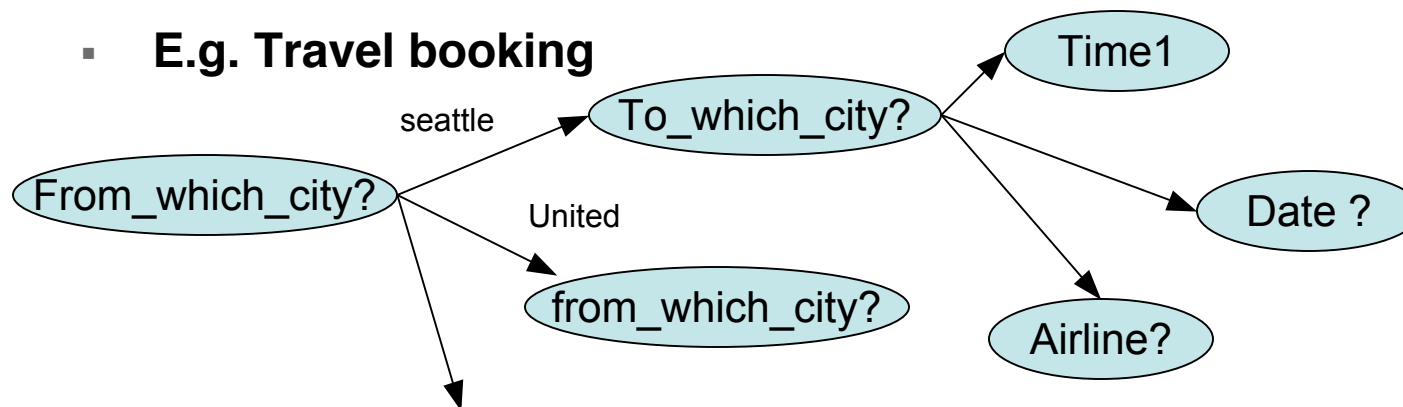
- **Adjacent turns by different speakers form a simple structure.**

First part	Second part
Greeting	Greeting
Question	Answer
Proposal	Acceptance/Rejection

- **Second-part follows the first-part in a coherent and predictable way.**
- **Can be implemented at dialogue-act or surface text level.**
- **E.g. ELIZA, chat-bot systems**
 - AIML, simple pattern matching and text rewriting
- **NPCEditor is best suited for this type of dialogue modeling. (E.g. SGT Star, VHuman you'll be building)**
- **Drawback : Does not model long-distance dependencies**

Finite State models (Pieraccini & Huerta, 2008)

- Define the complete voice user interface (VUI) as a FSM
- *Directed dialogue Or System-initiative*
- Designed for usability
- Useful for naïve users who don't know the task structure
- Commercial applications, IVR
- Drawback : complexity increases rapidly, very-restrictive dialogue flow for experts
- E.g. Travel booking



Form-filling models (Goddeau et al, 1996)

- Dialogue State is captured by a form structure
- Form is a set of <slot,value> pairs
- Useful for information exchange dialogues
- Allows for **mixed-initiative** (over-specification)
- Form Interpretation Algorithm (FIA) in VoiceXML
 - The goal is to fill all the slots.
 - Each input can fill multiple slots.
 - Each slot value can be elicited by a user designed prompt.
- **Drawback: Static form structure**

Slot	Value
from_city	Los Angeles
to_city	
departure_date	
Airline	United

S: Which city are you leaving from?

U: I'm on a united airlines flight from Los Angeles ..

VoiceXML example

```
<vxml version="2.0" lang="en">
<form>
  <field name="from_city">
    <prompt>Where do you want to fly from?</prompt>
    <option>Edinburgh</option>
    <option>New York</option>
    <option>London</option>
    ...
  </field>
  <field name="to_city">
    <prompt>
      Leaving from <value expr="from_city"/>,
      Where do you want to fly to?
    </prompt>
  </field>
  <block>
    <submit next="someURL" namelist="city travellers"/>
  </block>
</form>
</vxml>
```

Agenda Based (Xu & Rudnicky, 2000)

- Collaboratively generate the **product**, hierarchy of forms

- **Product**

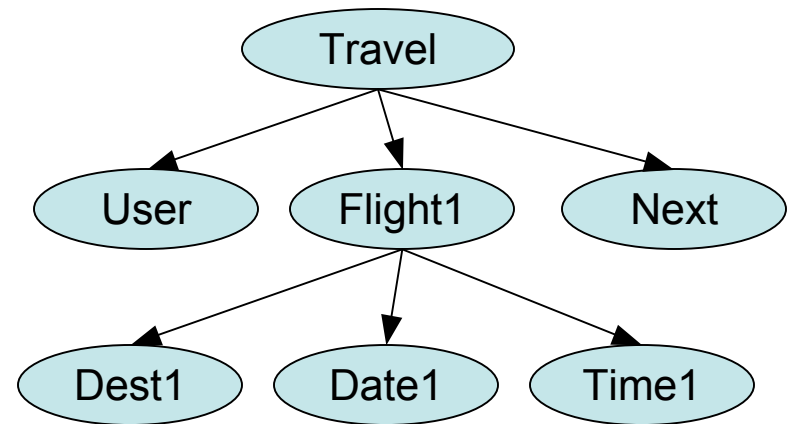
- Generalization of stack
- a tree structure of handlers (forms)
- Accessible by both dialogue participants

- **Agenda**

- Ordered list of handlers

- **Control Algorithm**

- Input passed to handlers on the agenda in order
- Output is prompts from handlers
- Mechanisms to implement topic shifts



Information-state update (Traum & Larsson, 2003)

- **Generalization of previous models (Comparing different dialogue theories)**
- **Informational components**
 - Shared (common ground, Last dialogue move, commitments, QUD)
 - Private (goals, beliefs, obligations, agenda, plan)
- **Formal Representations**
 - Records, propositions, DRSs
- **Dialogue moves**
 - Each dialogue move will update the information-state
 - Different dialogue act schemas
 - DAMSL (Core and Allen, 1997)
 - DIT++ (Bunt, 2006)
 - Your own custom designed for application
- **Update rules**
 - Preconditions (evaluated on the current information-state)
 - Effects (changes made to information state)
- **Update strategy**
 - Fire all rules that can be applied
 - Fire the rule with maximum utility
- **Toolkit released (TrindiKit, <http://www.ling.gu.se/projekt/trindi/trindikit/>)**

Information-state update (Poesio & Traum, 1998)

- **Social Commitments (public counterpart of private belief)**
- **Obligations (Traum & Allen, 1994)**
- **Conditionals (look-ahead rules for obligations and commitments)**

$$\begin{array}{l}
 \left[\begin{array}{l}
 G : PT-R \\
 CDU : \left[\begin{array}{l} C : PT-R \\ ID : DU-ID \end{array} \right] \\
 PDU : \left[\begin{array}{l} C : PT-R \\ ID : DU-ID \end{array} \right] \\
 UDU_s : List(DU-ID) \\
 INT : List(Action)
 \end{array} \right]
 \end{array}
 \quad
 PT-R : \left[\begin{array}{l}
 DH : List(Action) \\
 OBL : List(Action) \\
 SCP : List(Prop) \\
 COND : List(Action)
 \end{array} \right]$$

Update rules

Dialogue Act	Obligation	Commitments	Conditionals
S_1 Promise A	S_1 Achieve A		
S_1 Request A	S_2 address request: accept A or reject A		<i>if S_2 accepts A, S_2 obliged to achieve A</i>
S_1 YNQ whether P	S_2 Answer-if P		
S_1 Assert P		S_1 committed to S_2 that P	<i>if S_2 accepts P, S_2 committed to S_1 that P</i>

Agent-based (Perrault & Allen, 1980)

- **Agent defined in BDI (Belief, Desire, Intention) framework**
- **Speech-Acts are plan-operators**
 - Preconditions
 - Body
 - Effects
- **Can use Plan construction (for generating) and Plan inference (for understanding)**

- **E.g. Inform**

INFORM(s, h, P) – Illocutionary act

prec: $\text{KNOW}(s, P) \wedge \text{W}(s, \text{INFORM}(s, h, P))$

effect: $\text{KNOW}(h, P)$

body: $\text{B}(h, \text{W}(s, \text{KNOW}(h, P)))$

S.INFORM(s, h, P) – surface act

effect: $\text{B}(h, \text{W}(s, \text{KNOW}(h, P)))$

Interface or Interlocutor

- **Beware of the metaphor reflected by your dialogue system.**
- **Virtual Humans**
 - Reactive behavior (Traum & Allen, 1994 gives high priorities to addressing obligations)
 - Allow mixed-initiative
- **Embodied Characters**
 - Input modalities
 - Gaze – Addressee identification
 - Hand Gestures – Reference resolution
 - Head nods – acknowledgements
 - Output modalities
 - Hand gestures body postures – convey emotions
 - Gaze – addressee information, turn-taking acts (release-turn, keep-turn)



Virtual Human Spectrum



SGT Star
Question-Answering



Tactical Questioning
Question-Answering with
goals, compliance levels



SASO-ST/EN
Negotiation

complexity

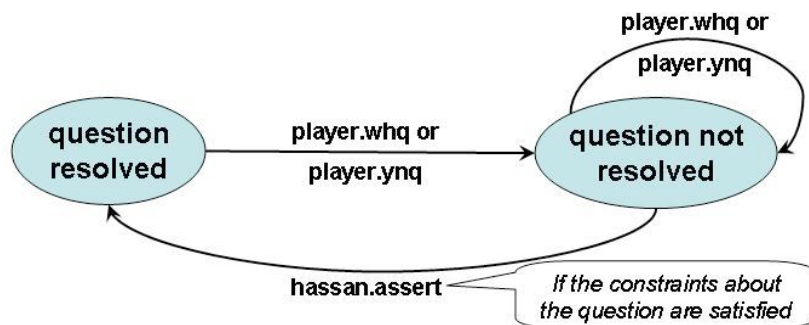
SGT Star (Artstein et al, 2008)

- **Question-Answering characters**
- **Maintains history of latest dialogue moves to avoid repetitions**
- **Off-topic response strategy**
 - **1. Didn't hear**
Could you throw that at me again?
 - **2. Didn't understand**
Sometimes I think you're just testing me.
 - **3.1 Move on**
You can get answers at GoArmy dot com.
 - **3.2 Prompt**
Hey why don't you ask me about my badges?



Tactical Questioning (Gandhe et al, 2008)

- Conversational games (Lewin, 2000)
- Tracking emotions, determining compliance level
- Answering questions only if certain constraints are satisfied
- Characters try to get what they want from the captain.



TacQ

<question resolved, offer not elicited>

player.whq *Ok, I'm trying to understand where the local taxation is coming from?*

<question not resolved, offer not elicited>

hassan.repeat-back
(grounding) *So you want to talk about the taxes.*

hassan.elicit-offer *I might tell you what you want if there was something in it for me.*

<question not resolved, offer elicited>

player.offer *We can offer you financial reward.*

<question not resolved, offer given> (constraint satisfied)

hassan.assert *Please understand, I collect taxes for my Imam. All in service to Allah.*

<question resolved, offer not elicited>

player.whq *And what is **his** name? (anaphora resolution)*

<question not resolved, offer not elicited>

hassan.elicit-offer ***My friend**, if people find out that I tell you this, it would be a problem for me. (style generation)*

Evaluation

- **Component by component evaluations**
- **Task-oriented systems**
 - Task success (# of slots correctly filled)
 - Task efficiency (# turns in a dialogue)
- **Non-task oriented systems**
 - No standard way
 - Pre and post questionnaires collecting subjective judgments
 - Turn-by-turn appropriateness ratings
 - Hand annotating dialogue corpus for correct information-state updates

References

- Sudeep Gandhe, David DeVault, Antonio Roque, Bilyana Martinovski, Ron Artstein, Anton Leuski, Jillian Gerten, and David Traum, [From Domain Specification to Virtual Humans: An integrated approach to authoring tactical questioning characters](#) to be presented at Interspeech 2008, September 2008.
- David Traum , Stacy Marsella , Jonathan Gratch , Jina Lee , and Arno Hartholt, [Multi-party, Multi-issue, Multi-strategy Negotiation for Multi-modal Virtual Agents](#) to appear in 8th International Conference on Intelligent Virtual Agents, September 2008.
- D. R. Traum, W. Swartout, J Gratch, S Marsella, "A Virtual Human Dialogue Model for Non-team Interaction", in Recent Trends in Discourse and Dialogue Springer, Laila Dybkjaer and Wolfgang Minker, Eds, pp. 45--67, 2008.
- David Traum and Staffan Larsson, [The Information State Approach to Dialogue Management](#) to appear in *Current and New Directions in Discourse and Dialogue*, Ed. Jan van Kuppevelt and Ronnie Smith, Kluwer, 2003.
- David R. Traum and James F. Allen [Discourse Obligations in Dialogue Processing](#) In *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics (ACL-94)*, pages 1--8, June 1994. [abstract](#)
- Emanuel A. Schegloff and Harvey Sacks: [Opening Up Closings](#) Semiotica, VIII, 4 (1973) 289-327.
- Xu, W. and Rudnicky, A. [Task-based dialog management using an agenda](#). ANLP/NAACL 2000 Workshop on Conversational Systems, May 2000, pp. 42-47.
- D. Goddeau, H. Meng, J. Polifroni, S. Seneff, and S. Busayapongchai. 1996. [A form-based dialogue manager for spoken language applications](#). In Proc. ICSLP, 1996 pp. 701--704.
- Perrault and Allen [A plan-based analysis of indirect speech acts](#). Computational Linguistics, 6:167-183, 1980
- [A formal model of conversational game theory](#) I. Lewin. In Fourth Workshop on the Semantics and Pragmatics of Dialogue: Gotalog 2000 , pages 115--122, 2000
- Massimo Poesio and David Traum, [Towards an Axiomatization of Dialogue Acts](#), In J. Hulstijn and A. Nijholt (eds.), *Proceedings of the Twente Workshop on the Formal Semantics and Pragmatics of Dialogues (13th Twente Workshop on Language Technology)*, Enschede, pp., 207--222, May 1998.
- Roberto Pieraccini and Juan M. Huerta, "Where Do We Go from Here? Research and Commercial Spoken Dialogue Systems". Recent Trends in Discourse and Dialogue, Text, Speech and Language Technology Volume 39, 2008.
- Ron Artstein, Jacob Cannon, Sudeep Gandhe, Jillian Gerten, Joe Henderer, Anton Leuski and David Traum. [Coherence of off-topic responses for a virtual character](#). To be presented at the 26th Army Science Conference, Orlando, Florida, December 2008



Questions

