

WTA 2015



FDTE - Desenvolvimento de Sistemas Digitais

Adaptalker: Um Framework para o Gerenciamento Adaptativo do Diálogo

D.A. Alfenas, C. E. Bogik, D. P. Shibata, R. P. da Silva, M. R. Pereira-Barretto

Gerenciamento de diálogo

- Sistemas de diálogo falado: projetados para o uso da voz como principal canal de interação com o usuário, através de linguagem natural
 - Inclui robôs sociáveis, sistemas de atendimento telefônico, etc.
 - Não inclui sistemas de ditado ou análise de voz (*speech analytics*)
- O gerenciador de diálogo é o processo central responsável pela tomada de decisão sobre *qual* a próxima ação a ser executada pelo sistema

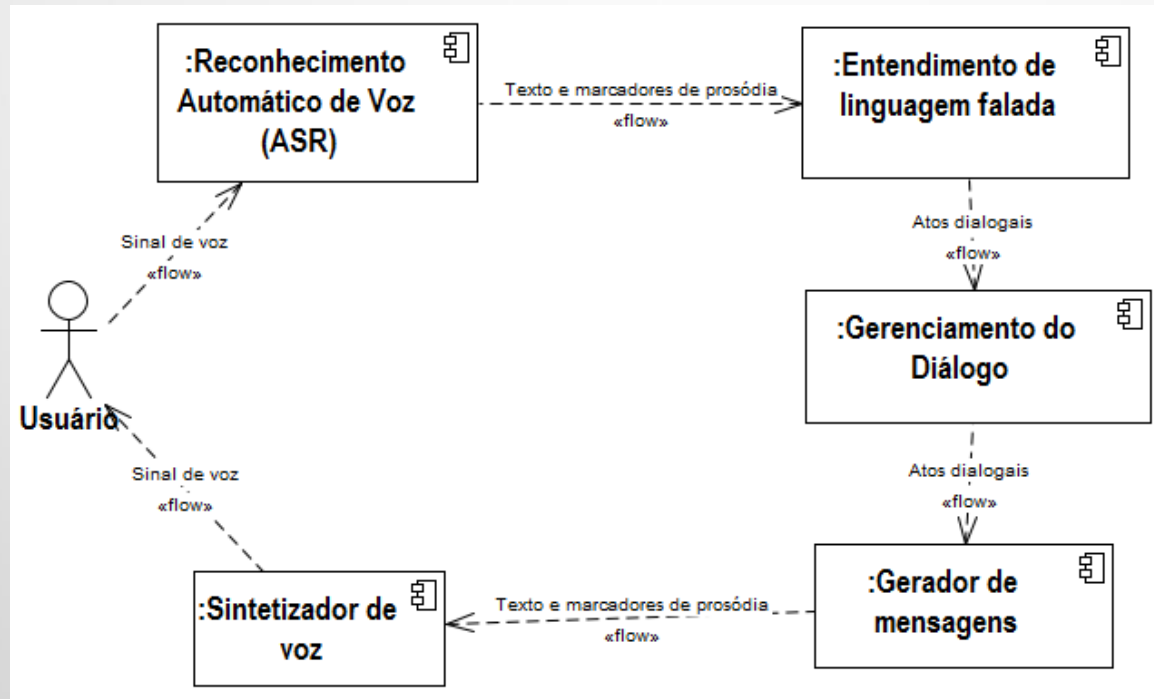
O diálogo

- Sistemas de diálogo se baseiam na definição de (Sacks, Schegloff, Jefferson 1974)
- Principais características:
 - A pessoa que fala muda ao menos uma vez
 - Composto por turnos
 - Existem mecanismos de reparo para lidar com situações de erro
- Os turnos são compostos por *unidades de construção de turno*
 - Frases
 - Palavras
 - *Segmento funcional*

Função comunicativa

- Tipo de intenção comunicativa. Exemplo:
 - Pergunta, resposta, *feedback*, elogio, cumprimento-inicial, cumprimento-resposta, etc.
- Ato dialogal
 - “Instância” de função comunicativa, isto é, uma ação de fato realizada por um participante para comunicar algo e atualizar o estado da informação nos ouvintes
 - Caracterizado por itens semânticos e relações funcionais com outros atos dialogais

Arquitetura típica de sistema de diálogo



Responsabilidades típicas de um gerenciador

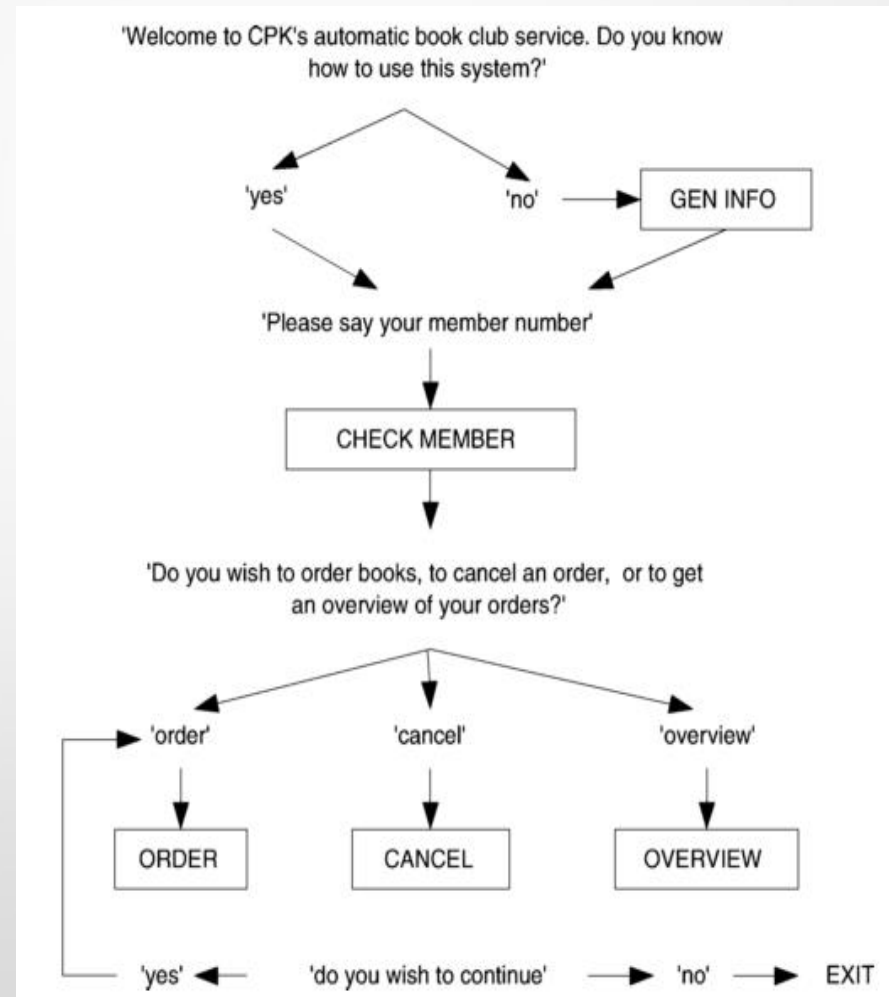
- Estimar o estado atual do diálogo
- Determinar quais as ações que o sistema deve realizar através de atos dialogais
- Para tanto, a *intenção* manifesta pelo usuário, capturada através de atos dialogais, é mais relevante do que as palavras exatas utilizadas
 - Importância do entendimento da linguagem falada a partir de *múltiplas hipóteses* sobre a ação do usuário

Técnicas de gerenciamento

- Máquinas de estados finitos
- Frames ou templates
- Atualização do estado da informação
- MDP ou POMDP
- Planejamento/Agenda

Baseado em máquina de estados finitos

- Popular comercialmente
- Facilita o desenvolvimento e testes
- Dificulta a modelagem de diálogos em que o usuário tem a iniciativa
- Dificulta a modelagem de diálogos com trocas sensíveis ao contexto, como o reparo



Adaptalker: Resumo

- Composto de duas camadas:
 - Filtro
 - DATD (Dispositivos Adaptativos de Tomada de Decisão):
 - Estimam o estado do diálogo (reconhecimento)
 - Fazem a tomada de decisão

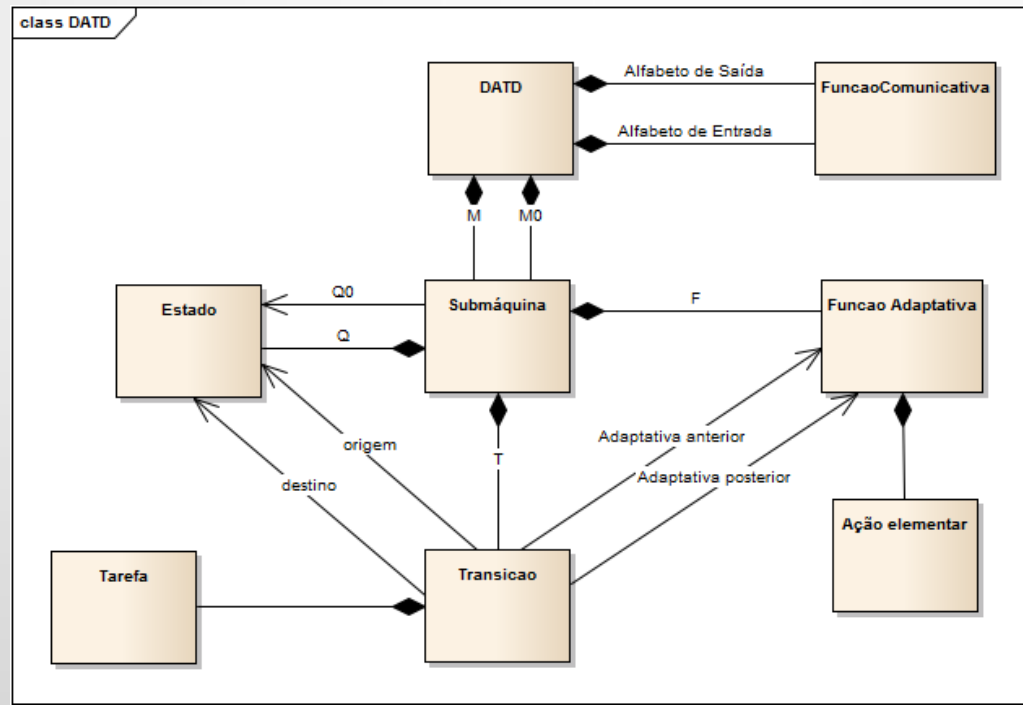
Adaptalker (cont.)

- Modelo de entrada e saída de atos dialogais baseado na norma ISO 24617-2, adaptado para suporte a múltiplas hipóteses através de eventos de atualização de hipóteses
- Na tomada de decisão, a adaptatividade é utilizada para:
 - Sensibilidade ao contexto, incluindo reparo
 - Mudanças de assunto decorrentes da iniciativa do usuário

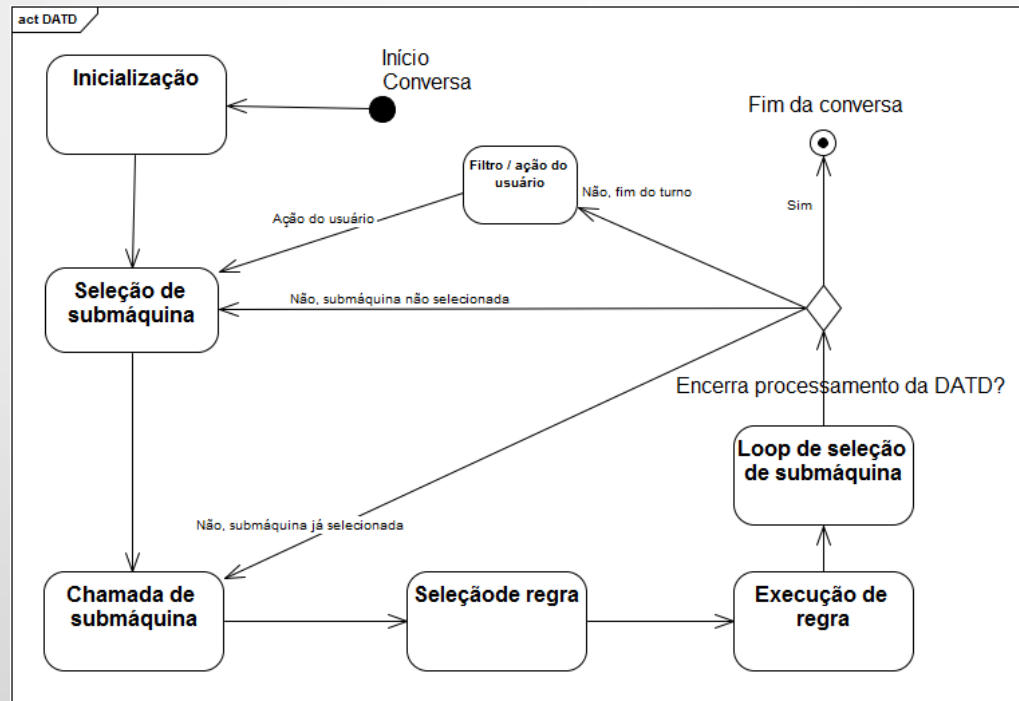
Filtro

- Envia erros gerados a partir do grau de confiança das múltiplas hipóteses e envia para o DATD, tais como:
 - Não entendimento total
 - Dúvida sobre a interpretação
 - Entendimento parcial
- Para cada hipótese, cria-se uma cópia diferente do DATD gerado ao final do turno anterior
 - Processa-se a hipótese
 - Calcula-se a probabilidade a posteriori
- Escolhe-se o DATD com maior probabilidade e descartam-se os demais

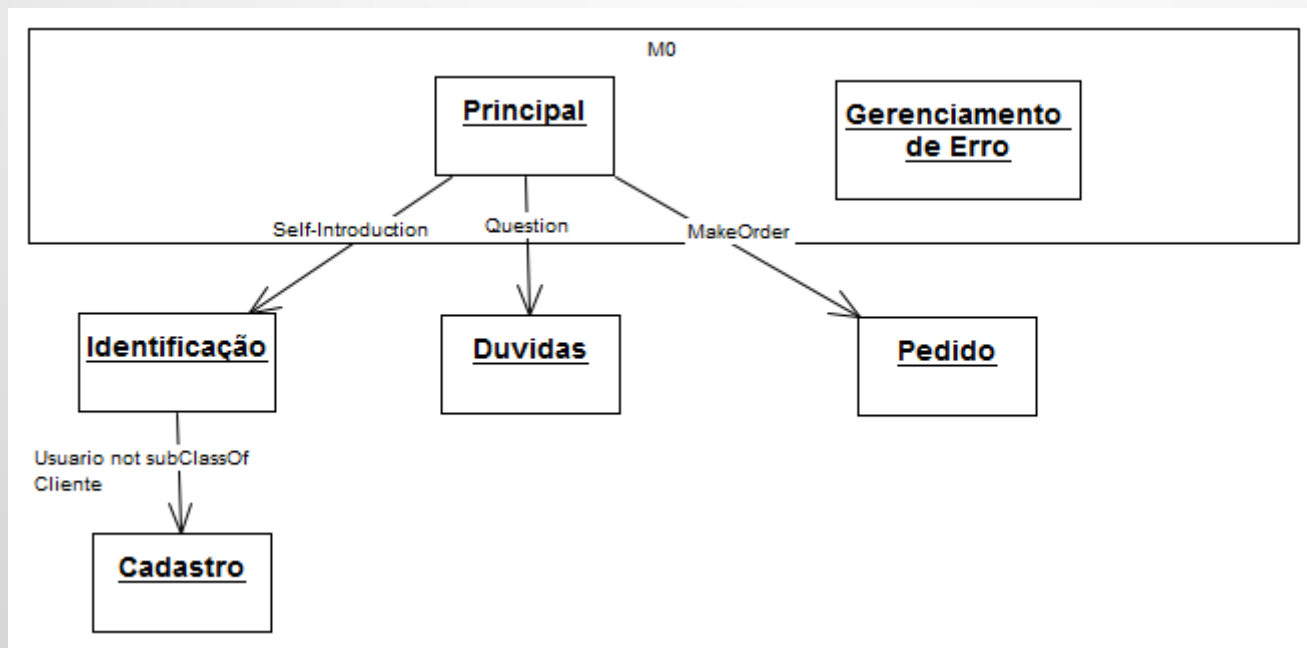
DATD: Visão estrutural de alto nível



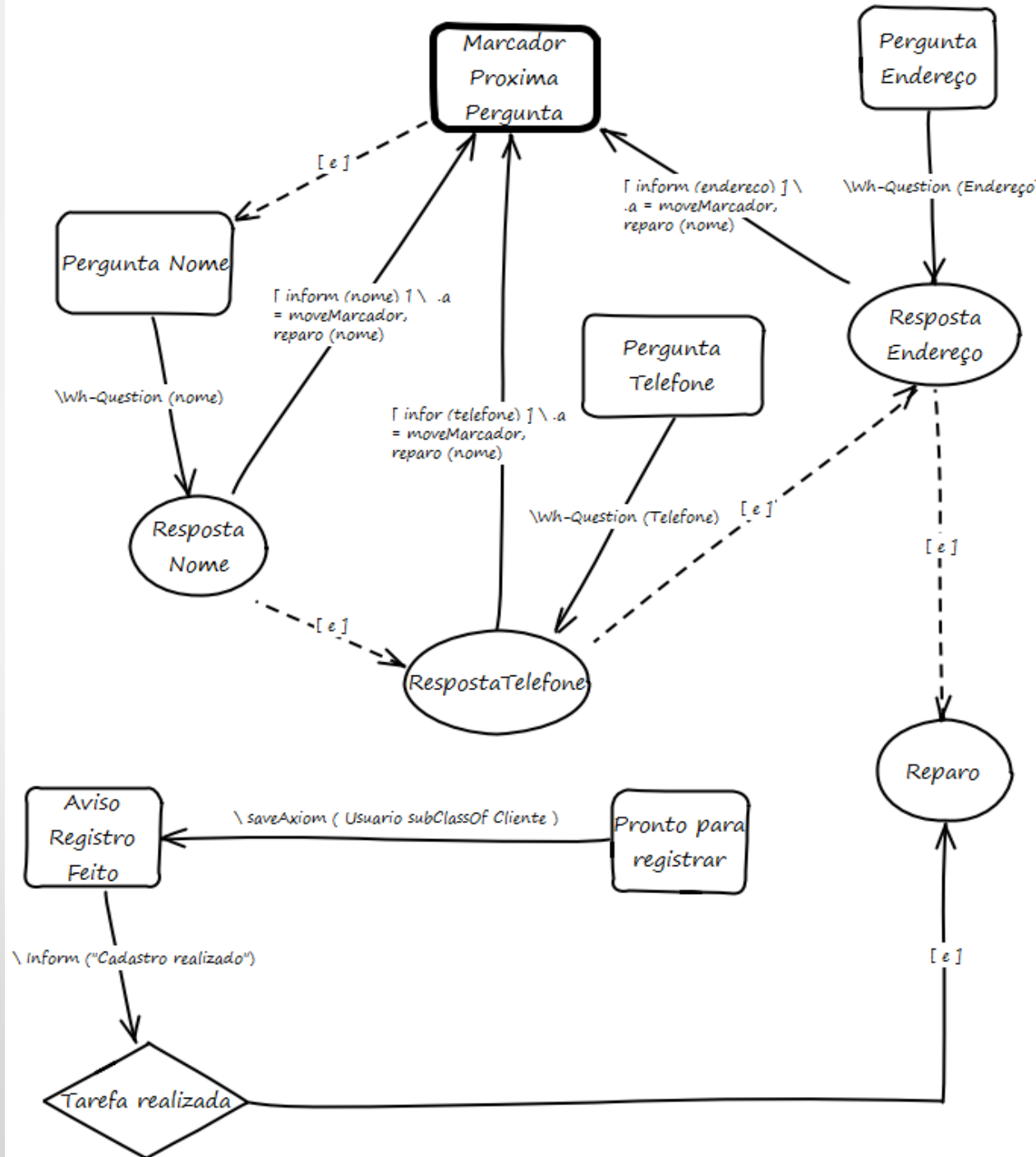
DATD: Algoritmo de execução



Exemplo de submáquinas



Submáquina de Cadastro



Comparação com critérios retirados da literatura

- Multimodalidade: Não testado
- Incremental: Não suportado
- Múltiplas hipóteses: Suportado
- Prosódia/emoção: Suportado
- Iniciativa mista: Suportado
- Sequenciamento: Suportado
- Facilidade de suporte ao reparo: Alta
- Facilidade de desenvolvimento: Alta

Comparação com o método não adaptativo original

- O uso de adaptatividade não retirou as vantagens do modelo original
- Benefícios da adaptatividade
 - Menos regras
 - Melhor modelagem de regras de iniciativa do usuário
 - Melhor modelagem de regras de reparo
 - Dependência de contexto
- Cuidados extras com regras adaptativas nos testes e desenvolvimento

Situação: Trabalho em Andamento

- Construção de framework conforme modelo construído em (Alfenas 2014) para permitir reuso
- Extração de características prosódicas para português brasileiro
- Construção de módulo entendimento de linguagem (SLU) falada operando sobre saídas de ASR comercialmente disponíveis
 - ASR testados com diversas qualidades de áudio
 - Comparação de diferentes técnicas

Trabalhos futuros

- Geradores de texto em linguagem natural
- Sintetizadores de voz: várias opções comerciais disponíveis
- Testes reais
 - Mágico de Oz

WTA 2015

Fim!

Interesse/dúvidas: d.alfenas@fdte.org.br

