Virtual DJs of KBOT:

Embedding Chatterbot Technology into Real-Time Media Streams

Doug DeGroot

LIACS, Leiden University

The Netherlands & Dallas

23 Feb 2003

degroot@acm.org

KBOT Description

- Full-fledged Internet Radio Station
- Run by Chatterbots and Info Agents
- Psychographically Targeted
- Designed to Support Groups of Listeners
- Supports both Entertainment and Informational Needs
- Interfaces to the Web, Email, and X-10
- Available 24 / 7

Research Goals:

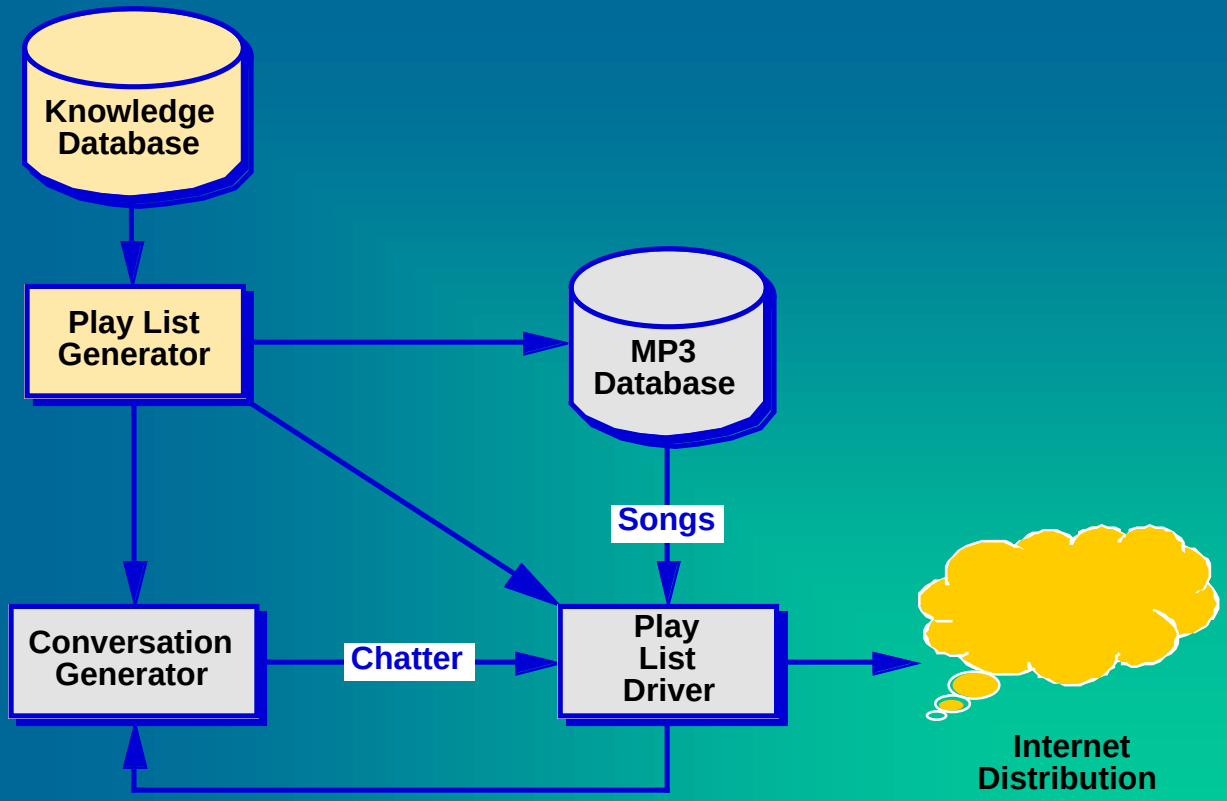
- 2-way chatterbots
  - Have two chatterbots carry on a single, *meaningful*, goal-directed conversation.
  - Limited to the “dueling banjos” model for now.
- Automated generation of “brain files”
  - Based on personality profiles & emotion states.
  - Customized for time-of-year, theme, etc.
- Chatter Grammars
  - Story-Telling-like Banter that is both Supportive & Emotive

Research Goals:

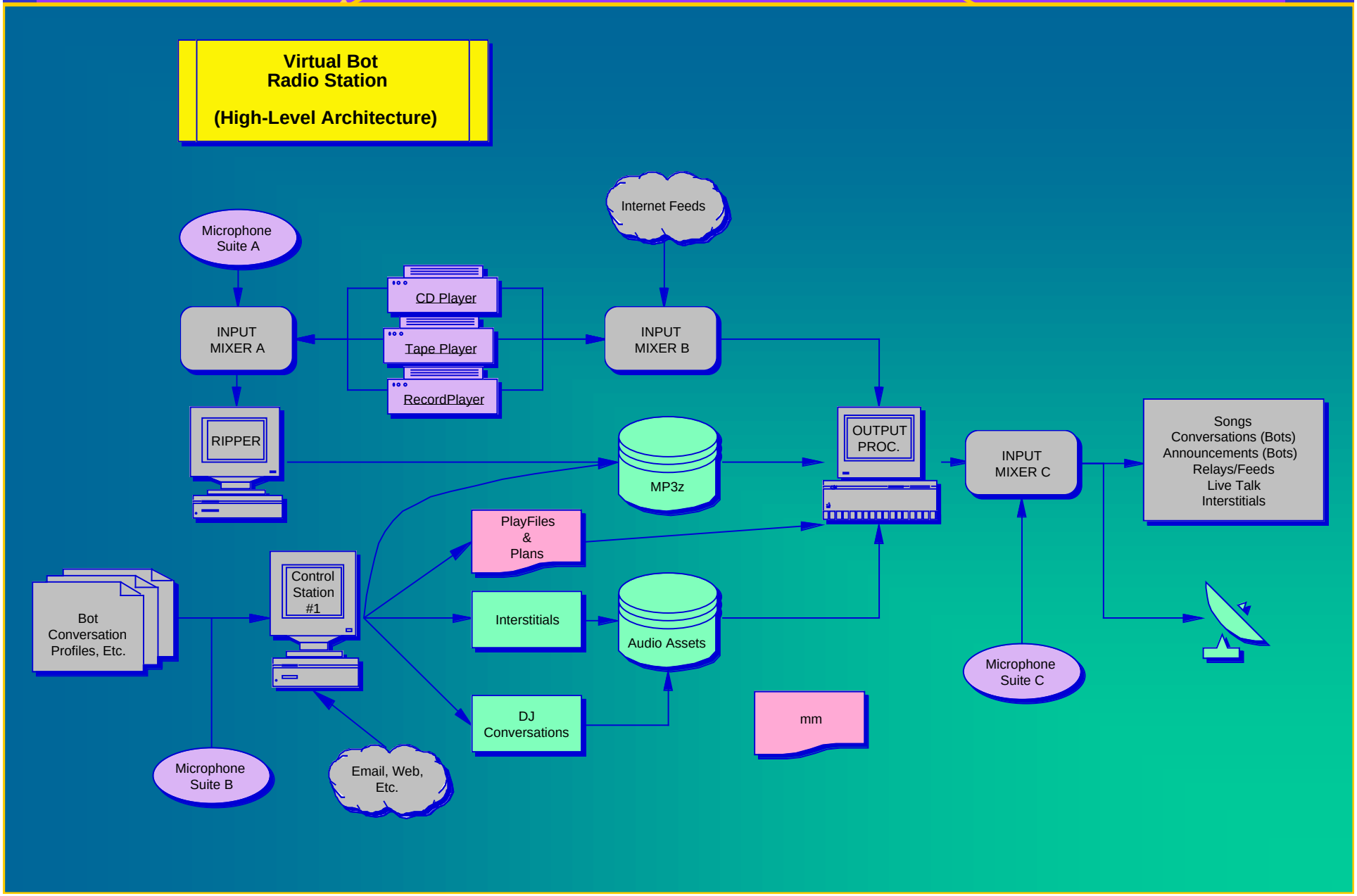
- Semi-automated text-to-speech markup
  - for emotional conversions
- Open-architecture for the Bot's Emotional Profiles
  - Users can collaborate & co-define.
  - Guest DJs.
- Achieving Suspension of Disbelief
  - When, how, and why?
- Goal-driven, automated "sound tweaking" for moods, branding, etc.

Top-Level Architecture of the Station:

- **Main Repositories:**
  - Song Libraries
  - Bot Brains
  - Bot Personalities
  - Play-list Segments
  - Bot-specific Knowledge
  - Personality-specific Knowledge
- **Main Processes**
  - Broadcaster
  - Play-list Generator
  - Bot Conversation Generator
- **Back-end Processes**
  - User-request Processor



Physical Layout of the Station:



Who's Who?

- **The DJ Bots**

  Multiple personalities, one or two at a time.

- **Weather Bot**

  Typical weather reporting functions (non-interactive in V1)

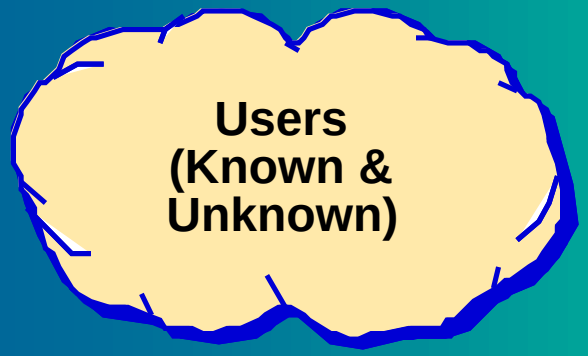
- **Email Bot**

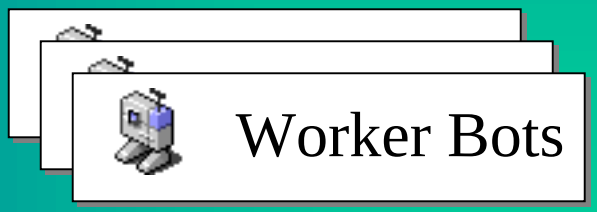
  Processes email requests. Can read or pass on to the DJs.

- **Producer Bot**

  Used by DJs to blame or explain (e.g., "Well. our producer tells me that …")

- **Worker Bots**





Key Aspects of the Design

- No Humans in the Loop
- Fully Real-Time, Dynamic, Live-Sounding
- Targeted to Specific User Groups
  - University / Department, family, club, etc.
- Extensible by Group Members
- Focus on Behavior, not Intelligence
- Suspension-of-Disbelief vs. Turing Test
- Both entertaining and helpful
  - to the group

Key Elements of the System

- **Text-to-Speech System**
  - Any COTS system that supports annotation
    - Emphasis, pitch, duration,etc.
    - Multiple Voices required
  - Our higher-level emotional “macro” markups
- **Conversation Generator**
  - Based on AIML syntax (cf. Alice Bot)
  - Uses Conversation Threads, Grammars, Personality Profiles, and “Job Lists” of the day

Key Elements of the System

- **Bots may be aware of external events & conditions:**
  - Who's listening (or logged in, at least)
  - Time, calendar, weather, station themes, ...
- **User's may be able to add value to the broadcast:**
  - Playlists - via personal listener profiles.
  - Bot Conversations - listener requests, "love notes", etc.
  - Extensible knowledge base for the Songs.

Time Objects - Domains & Themes:

- Seasonal Themes
  - Semesters
  - Summer
  - Skiing Season
  - Christmas
- Weekly Themes
  - Vacation
  - Spring Break
  - Christmas
- Daily Themes
  - Holidays
  - Two-for Tuesdays
  - TGIFs
  - Rainy Days
  - Weekends
- Hourly Themes
  - Bone-a-thons
  - Oldie-Goldie Hour

Themes are a key component of the conversation:

- Seasonal Themes
  - Semesters
  - Summer
  - Skiing Season
  - Christmas
- Weekly Themes
  - Vacation
  - Spring Break
  - Christmas
- Daily Themes
  - Holidays
  - Two-for Tuesdays
  - TGIFs
  - Weather
  - Weekends
- Hourly Themes
  - Bone-a-thons
  - Oldie-Goldie Hour

Her: Hope you're all still enjoying your summer vacation here on the one and only Rock-to-the-Bone KBOT zone.

Him: Well, here we are dog humans, celebrating the end of another week of final exams. So for all you goofers out there, we're gonna roll some ZZ Top for your enjoyment.

Her: We just heard a couple of top tunes from ZZ Top on this fabulous Two-for-Tuesday, coming at you all day, all the way. So now...

Him: OK dogs, settle back and relax, wherever you are, cause we have a couple of rainy day classics coming at you to help you forget about all this freaking RAIN!
Her: That's right, Joe Bot! So Hold the Doors for me, won't you?
Him: OK, here's Rainy Day Woman and then Riders on the Storm - only on the Bone with KBOT.
Her: Alrighty, then!

Him: Well, Humies, we're in the middle of an All-Classics Bone-a-Thon, coming up here on the Christmas Holidays.

What the Bots know:

- **The Station Model:**
  - The other bots at the station
  - Who's who & who does what.
- **Their current work model:**
  - Are they alone in the booth, or teaming up?
  - The producer and/or weatherBot, for example?
- **Who's online at the moment (and their profiles).**
- **Time elements:**
  - Time-based themes.
  - Where they are in the current time block (time left till done, etc.)
  - Clock time.
- **The online CDDB information (see later notes).**
- **Their own likes & dislikes concerning songs, groups, each other.**
- **Blah, blah, blah.**

The CDDB Album Data Fields:

- **Album Title:**
  - Includes sort information so “The Wall” can be sorted under ‘W’
- **Album Artist:**
  - Includes sort information so “Dave Matthews Band” can sort with “Matthews”
- **Record Label:**
  - The label or publisher of the CD
- **Year:**
  - The year the CD was recorded or published
- **Genre:**
  - Every album can have both a primary and a secondary genre (see below)
- **Compilation:**
  - Flag set for soundtracks, samplers, etc. Indicates whether tracks have different artists
- **Number/Total Set:**
  - Can identify a CD as a member of a box sets
- **Language:**
  - Used to help display in appropriate character set
- **Region:**
  - To identify where the CD was released
- **Certifier:**
  - Authorized party (artist or label) who has certified the accuracy of the data
- **ISRC:**
  - The International Standard Recording Code number for the CD.
- **Notes:**
  - General notes such as dedications, "recorded live in Tokyo", etc.

- **Gracenote CDDB tools.**

The CDDB Track Data Fields:

- **Track Title:**
  - Also includes sort information
- **Track Artist:**
  - Vital for compilations, such as soundtracks or samplers
- **Record Label:**
  - May be different from track to track for compilations
- **Year:**
  - E.g. in an anthology ("Greatest Hits"), may be different from track to track
- **Beats Per Minute:**
  - Used for DJ syncing
- **Credits:**
  - E.g. Guest Musicians, Songwriter, etc. (see below)
- **Genre:**
  - Every track can have both a primary and a secondary genre (see below)
- **Notes:**
  - General track notes such as “Recorded in St. Croix”, "Bonus Track", etc.

- **Credits:**
  - Can be entered for entire album, or applied to individual tracks or segments (150 fields)
- **Credit Name:**
  - Can be person, company, or place such as recording location
- **Credit Role:**
  - Musical instrument, composer, songwriter, producer, mixing, recording location, etc.
- **Credit Notes:**
  - E.g. to specify unusual instruments, "appears courtesy of ...", etc.
- **Genres:**
  - Can be entered for entire album or applied to individual tracks
- **Metagenres (20):**
  - General classification. E.g. Rock; Classical; New Age; Jazz
- **Subgenres (251):**
  - More specific style. E.g. Goth Punk, Ska; Baroque, Choral; Ambient; Bebop, Ragtime
- **Segments:**
  - Can be used to identify classical pieces (which can cross track boundaries) For musical pieces (groups of track movements) or to identify famous or critical sections Each segment can have its own name, notes, and credits

The KBOT Extra Data Fields:

KBOT Extra Track Data

- Track Intros:
  - <Sit-Backs>
  - <Call Outs>
  - ...
- Track Wrap-Ups:
  - <Howl Types>
  - <Oh Yeahs>
  - ...
- Track Theme Groups & Interludes
  - <Girls' Nite out!>
    - <Interlude 1>
    - <Interlude 2>
  - <Mellow Yellow>
    - <Interlude 1>
    - <Interlude 2>
    - ...
- P-A-D Emotional Points
  - Per theme type
- Listener Notes
- Suggested Threadings:
  - <Next Song, Score, PAD, Theme>
  - <Next Group, Score, PAD, Theme>
- Blah, blah, blah.

KBOT Extra Album Data

- Same Elements as Track Data
- Suggested Interstitials
- Suggested Threadings:
  - <Next Album, Score, PAD, Theme>
- Pointer list for graphics, lyrics, etc. for web site integration.
- Blah, blah, blah.

Text Generation Technologies:

- **A derivative of JanusNodes & TextDNA:**
  - Text-based Delimited Natural-Language Algorithms
  - Derived from earlier versions of McPoet.
  - Composed of
    - TextDemons
    - TextDNA
  - Specifies the manner in which words and sentences can be strung together to make a conversation.
- **Reverse "Augmented Transition Networks"**
- **Jargonization techniques**
  - Similar to the "Encheferizer" techniques by J. Hagerman
  - Allows multiple personality "types" *(stereotypes?)*
    - Valley Girl, Hippy, Cowboy, Jiver, Lady Killer, etc.
    - Not to be confused with specific "personalities".
- **Text Abstraction techniques**
  - For preparing possible "talking points" for the DJs.
  - Extensive CDDB materials available, plus user annotations.

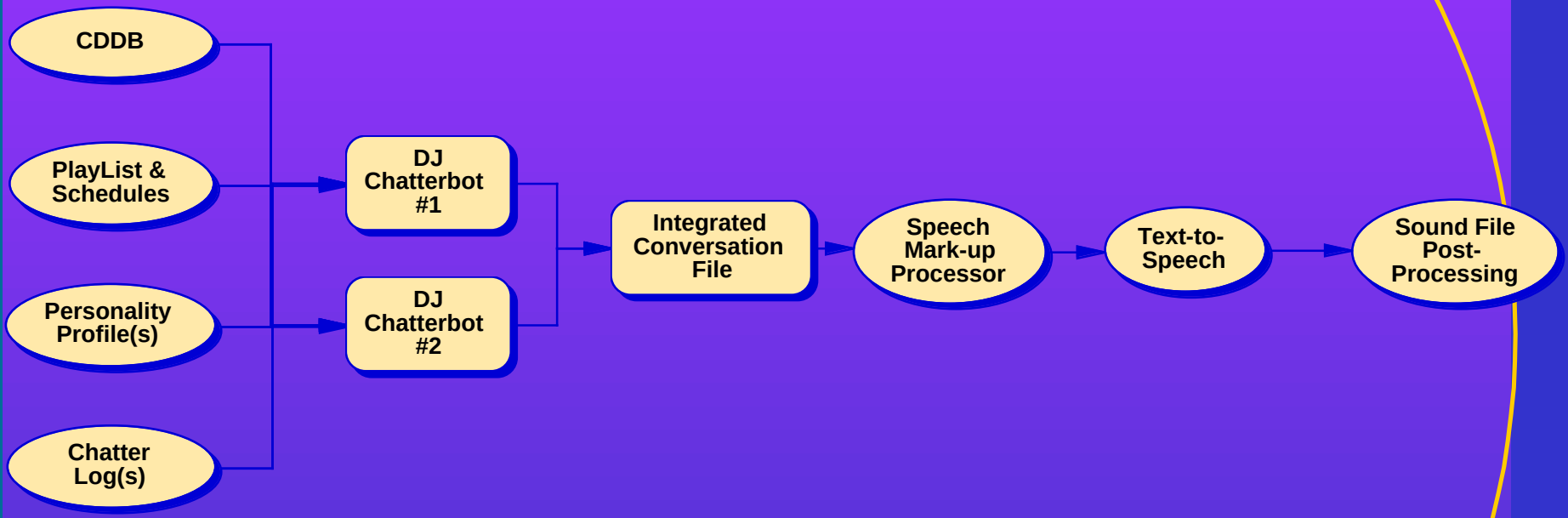
The PAD Emotional State Model :

- We implement dynamic “Emotional States” for each DJ.
- We are currently employing Mehrabian’s P-A-D Emotional State Model.
  - This model has both a theoretical rationale and experimental foundations.
  - Extensively documented, e.g., Mehrabian (1980, 1995b, 1997).
- The Model consists of three nearly independent dimensions
  - P: pleasure - displeasure
  - A: arousal - nonarousal
  - D: dominance - submissiveness
- PAD values can be used to
  - describe, compute, and measure emotional states (or feelings, affective conditions).
- Definitions:
  - pleasure - displeasure: distinguishes the positive-negative affective quality of emotional states
  - arousal - non-arousal: refers to a combination of physical activity and mental alertness
  - dominance - submissiveness: is defined in terms of control versus lack of control.

The PAD Emotional Scales :

- Specific emotions (terms describing emotions) can be visualized as points in a three-dimensional PAD emotion space.
- Alternatively, when the PAD scale scores are standardized, each emotion term can be described in terms of its values on the P-A-D axes.
- These sample ratings illustrate definitions of various emotion terms (when normalized from -1 to +1):
  - angry (-.51, .59, .25)
  - curious (.22, .62, -.01)
  - elated (.50, .42, .23)
  - inhibited (-.54, -.04, -.41)
  - puzzled (-.41, .48, -.33)
  - unconcerned (-.13, -.41, .08)
  - bored (-.65, -.62, -.33)
  - dignified (.55, .22, .61)
  - hungry (-.44, .14, -.21)
  - loved (.87, .54, -.18)
  - sleepy (.20, -.70, -.44)
  - violent (-.50, .62, .38).
- For example:
  - "Sleepy" consists of a moderately pleasant, extremely unaroused, and moderately submissive state.
  - ”Bored” is composed of highly unpleasant, highly unaroused, and moderately submissive components.

Generating Emotional Speech:



Meeting the Listener's Needs.

- Listeners can listen all day long.
- They may influence the playlist via personal profiles (even 1-on-1 playlists).
- They can have the DJs play songs for other listeners, along with dedications.
- Can expect the DJs to remind them about
  - Exams, Late trains, etc.
- Can co-architect the Bot personalities and databases.

Phased-In Deployment & Design:

- Phase 1:
  - Single commentary, hard-wired scripts.
  - Dynamic Emotional States, affects chatter.
- Phase 2:
  - 2-way autonomous chatter, shared goals.
  - Dynamic weather, emails, events, etc.
- Phase 3:
  - Multiple Personalities, dynamic conversation filtering.
  - User-extensible knowledge DBs.