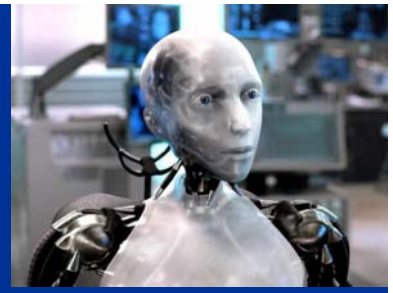


La Computación del Futuro

Agentes Inteligentes de Diálogo
¿Realidad o Ficción?

Andrés T. Hohendahl

Investigación y Docencia

Laboratorio de Estereología y Mecánica Inteligente
Facultad de Ingeniería, UBA.
IIBM (Instituto de Ingeniería Bio Médica)
Facultad de Ingeniería, UBA.

Empresas y Emprendimientos

CEO, www.PandoraBox.com.ar (software & services)
CEO, MoneyBox s.a. (mobile payment)
AndySoft (1982-1990)

Agentes de Diálogo

¿Qué Cosa Son?

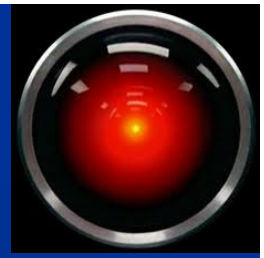
- Es un sistema que puede interactuar con humanos
  - Apto para fines Específicos y Determinados
  - Utilizará Tecnología de Comunicación (teclado, voz, cámaras)
  - Responderá en forma amigable (imágenes, voz y texto)
  - Manejará otros Recursos Digitales (archivos, fotos, mails)
- Estará Diseñado para
  - Entenderá Consultas en un Contexto Acotado
  - Brindará Ayuda sobre como dialogar con él.
- No será apto para
  - Entender cualquier cosa en cualquier idioma
  - Comprender cosas mal formuladas o incoherentes
  - Conversar sobre la Inmortalidad del Cangrejo
  - Intuir, leer entre líneas, inventar lo que no se le dijo, etc.

Para esto Necesitarán:

Procesamiento de Lenguaje Natural + Inteligencia Artificial

Agentes de Diálogo

¿Qué Pueden Hacer?



- Responder Preguntas
  - Soporte de Productos Tecnológicos, entre otros
  - Evacuar Dudas sobre productos y servicios

- Brindar Acceso a Servicios
  - Realizar Transacciones de todo tipo
  - Obrar por cuenta nuestra ante alguien más
  - Realizar Búsquedas de Referencia Virtual (Bibliotecas)

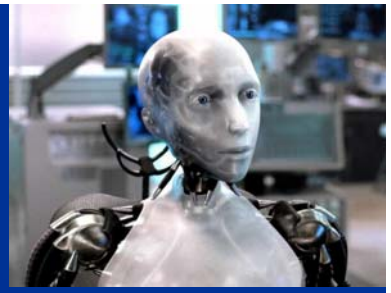
- Dialogar
  - Explicarnos cosas y ayudarnos con las dudas
  - Darnos acceso fácil y claro a Información
- Gestionar
  - Podrán Gestionar y Devolvernos resultados
  - Tendrán manejo de datos (archivos, etc.)

→ Para esto necesitan tener incorporado:

Procesamiento de Lenguaje Natural + Inteligencia Artificial

Inteligencia Artificial

¿En Que Consiste?



- Es el conjunto de los técnicas para hallar soluciones 'satisfactorias' a problemas irresolubles en forma directa o por fuerza bruta.
  - Algoritmos de Aprendizaje Automático

    Neural Nets (ANN), SVM, HMM (Markov Models), Simulated Annealing, Genetic Programming, Automated Learning, Decision trees (ID3, C4.5), Conditional Random Fields, Clustering, Agents, Dynamic Programming, etc.
  - Lenguaje Natural
    - Procesamiento de Lenguaje Natural (PLN-es / NLP-en)

      Tokenizers, Chunkers, Grammatical Parsers, Information Extraction, Speech Recognition (ASR), Analyzers, Prediction, Classifiers, etc.
    - Generación de Lenguaje Natural (GLN-es / NLG-en)

      Planning, Text to Speech (TTS), Prosodic Modeling, etc.
  - Vision Computarizada

    Lectura de Texto (OCR), Fingerprints/Face Detection, Biometrics, etc.
  - Robótica (Maquinarias)

    Movement-Planning, Minimal Path, Dynamic/Adaptive Controls,

El Lenguaje Natural

Es Eficiente y Compacto

Lógicamente: Los Humanos lo inventamos para comunicarnos...

→ ...Pero la Computación, reinventó eso... lamentablemente!

Veamos que tan complicado es hoy, con la herramienta más difundida!

Cómo escribir algo como.: 10^3 kg usando MS-office (aquí mismo)

1. Tipear '10'
2. Hallar la etiqueta de menú: Formato, Click!
3. Escoger el sub-menú: Fuente, Click!
4. Buscar un Bloque que diga Efectos, Hallar Superíndice
5. Apuntar a un cuadradito (al lado), Click!
6. Ir al borde de arriba con la cruz, Click!
7. Tipear '3' ....listo!.. No! (pues quedó todo chiquitito y arriba!)
8. Hay que rehacer todo y des-clickear ese casillero (pasos 2..5)
9. Tipear ' kg'

Además PREVIAMENTE DEBEN SABER que ESO QUE DESEABAN está detras de:

Formato>Fuente>Efectos/Superíndice>Tildar (poner el '3')

Formato>Fuente>Efectos/Superíndice>des-Tildar

Y con Texto Natural, como sería?

Tipear '10 al cubo kg' o tal vez '10 a la 3 kg' o simplemente 'mil kilos'

y... ¿Qué es más sencillo? (suponiendo que lo debemos aprender)

El Poder del Lenguaje Natural

La Espada de Grayskull

Se imaginan estar frente a un SISTEMA NUEVO del cual Ustedes: no saben cómo hacer algo que necesitan?

Que Harían?

1. Leer o lo buscan en un manual? (eso no se los creo...)
   - ESE Manual estará escrito en Lenguaje Natural
2. Se Capacitan (hacen que alguien se los explique.. más fácil)
   - ESE Alguien se lo explicará en Lenguaje Natural
3. Prueban de todo hasta que lo sacan (no es muy recomendable..)
   - Pero perderán mucho tiempo y pueden hacer macanas!

Pero Uds. Tienen el poder de Grayskull !

Pues Uds. SABEN LO QUE QUIEREN HACER

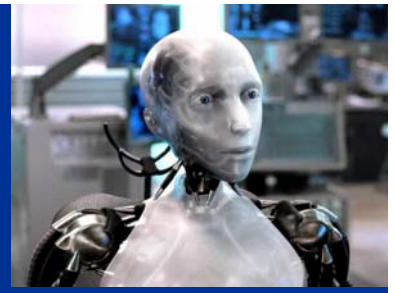
En otras palabras: lo pueden indicar o preguntar a alguien más!
→ O sea podrían indicarle al SISTEMA, exactamente qué desean hacer..
→ Y el SISTEMA (si es que los entiende) Simplemente lo hará,

→ esto es GENIAL!

→ Esto es precisamente el poder oculto del Lenguaje Natural!

Agentes de Diálogo

¿Cómo lo hacen?

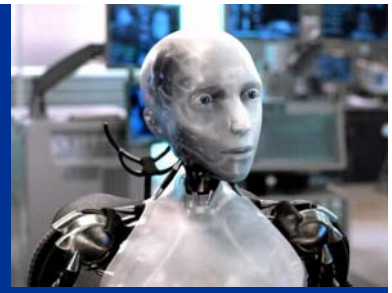


1. Entender
   - Debe poder Interpretar Preguntas para Responder
     - Si no está claro debe poder pedir Aclaraciones
   - Debe Desambiguar y Resolver Anáfora y Deixis
     - De allí extraerá la Pragmática
2. Resolver y Decidir
   - Obviamente debe poder resolver lo solicitado
     - Si está en el contexto permitido y para el cual fue diseñado
     - Accederá al los datos y/o resolverá la gestión encomendada
     - Obtendrá la información solicitada, de existir.
3. Responder, Comentar o Preguntar
   - Debe poder formular la respuesta
     - Generación de Lenguaje Natural (GLN)
     - Estilo Literario: Planning, Algoritmos
   - Debe manejar los Turnos de Diálogo

➔ Se Necesita un: **Motor de Diálogo**

Entender Texto

Humanizando la PC



Hooolaah

“tal vez quiso escribir ‘hola’..”

Reparación de Ortografía

vahyema

“esta suena como ballena..”

Similitud Fonética

Hoy a las 2 y media am

“puso una fecha/hora..”

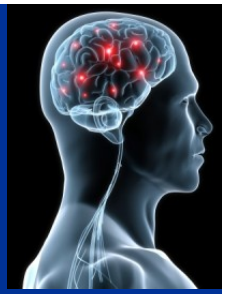
Servidor Léxico

Hohendahl

“esto parece alemán..”

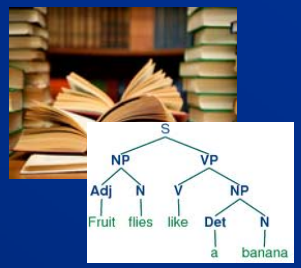
Estimador de Idioma

Secuencia de Procesos
Para Entender un Texto



- Analizador Morfológico (Lematización Robusta)
  - Si es palabra: estimar el idioma, o ver si es ruido
  - Reconocer: Números, Siglas y Palabras Flexionadas
  - Diferenciar y Clasificar: Términos Desconocidos
    - → Detección + Corrección de Errores de Ortografía

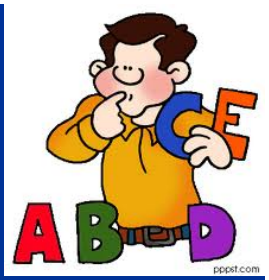
- Análisis Gramatical y Semántico
  - Reconocer Estructuras y Segmentos Complejos
    - Estadística / Autómatas / Gramáticas (CFG's, HRG)
    - Desambiguación (Gramatical y Semántica)



- Comprensión de Texto (Machine Understanding)
  - Análisis Pragmático
  - Resolución de Deixis, Anáfora y Coreferencia
  - Representación de Conocimiento (Modelos)
  - Análisis de "Sentido" (Common Sense)

Qué cosa quiso poner...?

Los Errores de Ortografía



Todo segmento de texto no es más que una triste chorrera de Letras...

Si no está en el diccionario → No se sabe nada a priori, por ej.:

K A V R H O M (7 letras)

~ Hay 35 letras diferentes c/signos diacríticos (acentos/diéresis/eñe)

- » Cambiar 1 letra ~ 7 x 34 ~ 234 intentos de búsqueda en diccionario
- » Cambiar 2 letras ~ 7 x 34 x 6 x 34 ~ 5 x 10^4 intentos
- » Cambiar 4 letras ~ 1.2 10^{12} intentos (sin contar eliminaciones ni inserciones)
- » Cambiar 7 letras ~ 9 x 10^{15} intentos (incluyendo duplicados)

Es un problema complejo y de orden combinatorio!

→ **NP Duro** = IRRESOLUBLE EN TIEMPOS RAZONABLES (POLINOMIALES)

Tampoco sabemos

- » El idioma (hoy todo es multilingual-mixto)
- » Si existe, en cuál diccionario está, si es un apellido o nombre propio.
- » Cuáles letras están falladas / faltan o sobran (debo probar cada letra)
- » Si se Invirtieron algunas letras y cuáles

→ No sabemos si se puede reparar en un tiempo/costo razonable

→ Si Hallamos varias alternativas (es probable)

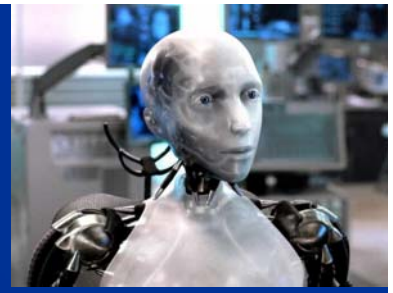
→ No sabremos jamás cual de ellas puede ser la más apropiada

→ Pero: un Humano resuelve esto en forma intuitiva y sin pensar mucho!

¿ Será **CABRÓN** ? (6 letras: 4 cambios + 1 eliminación)

I.A.+ Lingüística

Horrores de Ortografía

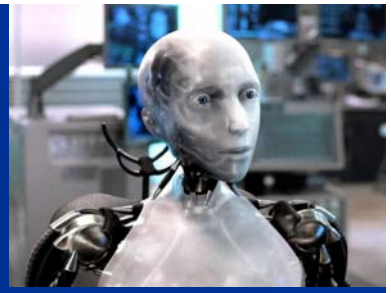


Análisis Morfológico y Lematización Robusta

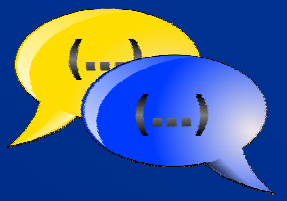
vimo hoi ezpe kavrom kon eza vayema ozpitalaria ke nempekapo salu2 klhdrdzkuio

vimo vino_VPIS3SM<*venir>~0.98 p:0.00000019 | vimo_Xes<*vimo>~0.83 p:0.036
vino_VIIS3SM<*venir>~0.98 p:0.00000019 | vino_NCMS~0.98 p:1
hoi hoy_RT~0.99 p:0.0000048 | ohm_NCMS0h~0.95 p:0.0000048 | Ho_NPMS~0.55 p:0.9 |
hot_AQ0MS~0.89 p:0.0000048 | hoi_Xen<*hoi>~0.84 p:0.9
ezpe éste_PD0MS~0.98 p:0.03 | este_DD0MS~0.98 p:0.9 | este_NCMS~0.98 p:0.0075 |
este_PD0MS~0.98 p:0.03 | Xes<*ezpe>~0.93 p:0.0075
kavrom cabrón_AQ0MS~0.96 p:0.76 | cabrón_NCMS~0.96 p:0.23 | Xde<*kavrom>~0.84 p:0.076
kon_con_SPSMS<%mode>~1 p:0.08 | kon_Xde<*kon>~0.9]~0,997
eza esa_PD0FS<*eso> p:0.076 | esa_DD0FS<*ese> p:0.9 | Xes<*eza>~0.93 p:0.076
vayema [ballena_NCFS0z~0.99 p:0.19 | ballena_AQ0FSL<*balleno>~0.99 p:0.62 |
ballena_NCFS0L<*balleno>~0.99 p:0.19 | vayema_Xes<*vayema>~0.88 p:0.062
ozpitalaria [hospitalaria_AQ0FS<*hospitalario>~1 p:0.039 | ozpitalaria_Xes<*ozpitalaria>~0.88]~0,995
ke _que_CS~0.99 p:0.011 | qué_PE0NS~0.99 p:0.032 | qué_DE0CN~0.94 p:0.000017 |
qué_PT0CNN~0.94 p:0.0015 | que_PR0CNN~0.94 p:0.009 | Xes<*ke>~0.55
nempekapo mentecato_AQ0MS~0.95 p:0.76 | mentecato_NCMS~0.95 p:0.23 |
Xen<*nempekapo>~0.789 p:0.076
salu2 saludos_NCMP<*saludo> | p:0.09 Ka<*salu2> p:0.9
klhdrdzkuio _B(BadWord)~0,875

Comportamiento
Resolver que Hacer

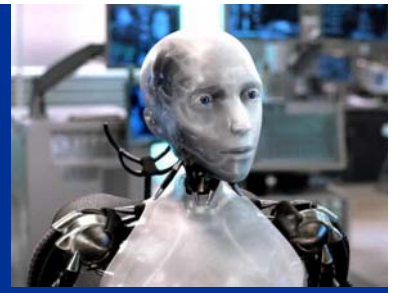


- Recetas del motor de Diálogo
  - DDL (Dialog Definition Language)
    - Entender (Pragmática)
      - Patrones Sofisticados (Parsing/etc.)
    - Operar con 'Palabras' y 'Conceptos'
      - Operaciones, Funciones, Coreferencia y Anáfora
      - Razonar (Motor de Inferencias)
    - Responder y/o Actuar en Consecuencia
      - Manejar Eventos, teniendo en cuenta el tiempo
    - Tomar Decisiones
      - Mediante Lógica y Deducción
    - Responder claramente al usuario
      - Generar Lenguaje Natural

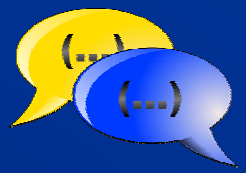


- Canales de Interacción
  - IRC (Protocolos de Chat y Mensajería)
    - MSN, ICQ, AIM, XMPP, GoogleTalk, SMS, XML, etc.

Comportamiento
¿Cómo Definirlo?



- DDL (Dialog Definition Language)
  - No es: Otro lenguaje de computación!
    - Está pensado para ser usado por humanos con formación normal
      no sólo por expertos en informática!

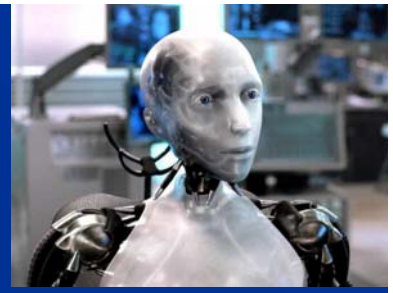


Opera con 'Palabras' y 'Conceptos' como lo haría un humano
- Ej. de Operaciones: + y *
  dos + tres → seis (número 6)
  dos * tres → llueve?
  MCMXI y veinte → 1931 (número)
  Karina + Andrea → Karina y Andrea
  gato + perro + gato → dos gatos y un perro
  N * m → joule (unidad física)
  décimo cuarto + uno → 15º (número ordinal)
- Funciones
  Similitud Fonética, Palabra + Importante/Relevante,
  Sinónimos, Hiperonimias, Holonimias, Paronimias
- Lógica y Toma de Decisiones
  Programación: Acceso a Bases de Datos y Servicios
- Cómo obrar si tal o cual cosa
  Respuestas y acciones ante Eventos

- Es Eficiente, Veloz y compacto
  Portátil y Compilado a MSIL y luego con JIT
  Tiene un compilador de verdad!, no es un script!
  Corre bajo .NET 2.0+ y Mono

Motor de Diálogo I

Imitando el Pensamiento



- Ejemplo de Eventos
  - Text Input: - 'hooolah komo estas!'
  - Corrección de errores: - hola cómo estás !
  - Análisis Morfológico-Gramatical y etiquetado:
    - greet[hola_I] ask[cómo_PT0MS estás_VMIP2S] emo[!_Fat]
  - Análisis Pragmático:
    - [time:initial] + [type:salutation] + [emotion:increase]
    - +Contexto/Expectativas/Memoria/Lógica
  - Answer: - Bien! ¿en que te puedo ayudar?

- Representación Interna del Mundo

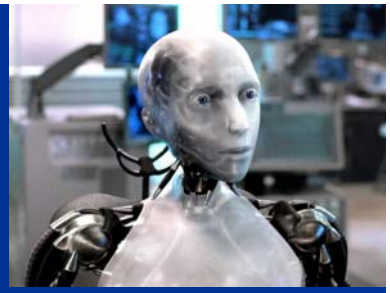
  Usuario
  - nombre/entorno/dudas/gustos/amigos/confianza/relatos/pertenencias/etc.

  Agente
  - datos/conocimientos/capacidades/conocidos/acceso a servicios/etc.

- Veamos una conversación SIMULADA (no real, pero sí: posible)
  - El Agente (bot) conversa con un humano (usuario), quien le relata una serie de cosas, el Agente lleva el diálogo y resuelve la conversación, en forma similar a un interlocutor humano, deduciendo relaciones lógicas y reconfirmando las que no son tan obvias.
  - Este tipo de diálogo es realizable con programación en DDL, usando los actuales motores de inferencia, anáfora y razonamiento que están siendo desarrollados.

Motor de Diálogo II

Anáforas y Relaciones



user Mi amigo se llama Carlos

bot Veo que conoces a Carlos!

Tiene una novia llamada María

También conces a María

Él tiene un android

Entiendo que Carlos tiene un android

Ella lo odia

Entiendo que María odia al android de Carlos

Él le puso Andri

Estimo que el android ése, se llama Andri

Si, y María tiene una amiga llamada Alicia

Que bueno es tener amigos!

A ella le gustan los celulares

Le gustan los androids a Alicia?

Si, Se lo dijo cuando ellas fueron al Shopping

A cual Shopping fueron María y Alicia?

A Unicenter, Su madre también odia los celulares

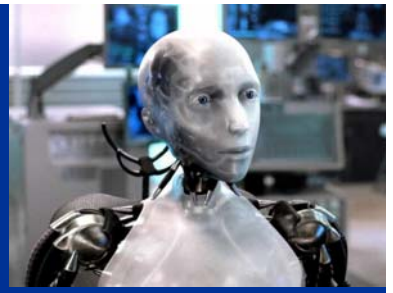
Cómo se llama la madre de María que odia los celulares?

Sandra

Entonces ella odia a Andri

El Pensamiento

Conocer y Razonar



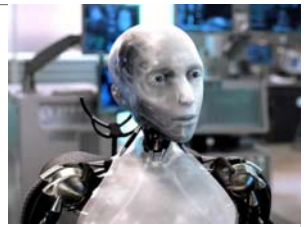
- Ontologías
  - Cómo Representar Conocimiento y Relaciones
    - OWL (web semántica), RDF, N3, Atom, etc.
    - Databases: Taxonomy, Biologcal, Chemical, Medical, etc.
      - » ChEBI / EFO / GO / SBO / SNOMED CT, etc.

EMBL-EBI

  - Palabras y Términos
    - Diccionarios y Thesauros
    - Semántica y la relación entre las cosas
      - SynSets (conjuntos de significados)
        - » WordNet: Inglés
        - » EuroWordNet: Español, Italiano, Alemán, Francés, etc.

- Razonamiento
  - Cómo Inferir nuevos datos en base a relaciones?
    - Incorporando el conocimiento del mundo (base)
    - Motor de Inferencias
      - » Reasoners (Prolog/Euler-algorithms/etc.)
      - » CYC (relaciones fácticas, en inglés solamente!)

Plataforma de Desarrollo
Agentes de Diálogo Inteligentes



SQL-DB storage

- Agent-Knowledge
- Anaphoric Memory
- WordNet
- Ontologic Knowledge
- Episodic Memory
- N3-OWL Reasoner
- User-Related Memory

Parser & Chunker GLR+

GLR Robust Linguistic Parser + Chunker

Agent-Interface Configurations

Converts DDL specification into MSIL ready fot the .NET RunTime Execution Unit

True-Compiler

DDL to **C#** Compiler

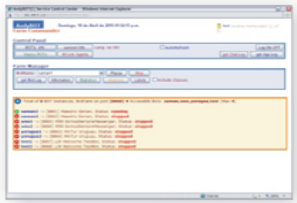
C# to **MSIL** Compiler

.NET Library Linker

***.DLL** Diagnostics

Debugger Benchmark

B A C K O F F I C E



App.: Bot-Farm Deployment

External Resources

Dictionary
Web-Wiki
Math.Libs
Symb.Math
Symb.Chemic
Scientific.Units
WebService
WMO Clima
Money-Rate

Agent Memory Reasoning & Relations

Grammatical & Semantic Functions

High-Availability Monitoring Service

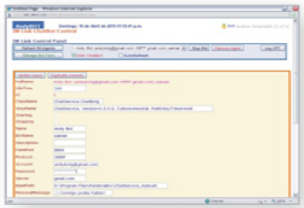
Library & External Functions

Runtime + IA. Dialog Engine

Control & Management

External Events

Response to Events



App.: IM Link ChatBot Control

Linguistic Resources

Lemmas
Locutions
Morpoho
Prefixes
Infixes
Suffixes
Semantic
WordNet
Statistics
Intl.Units
Numbers
Romanic
TimeDate
Money

Phrase Conformer + DAML Classifier

Default Knowledge Facts & FAQ´s

DDL Rules Execution

Provides Special Event/Interfaces for High-Demand Massive formats like Mobile e-mail & Robotics.

General Event Interface (Text/Actions/Files/Data)

Corrected +Tagged Text

Event Flow

Built-In Library: IA-Math-Lexic-Logic

package #N
package #3
package #2
package #1

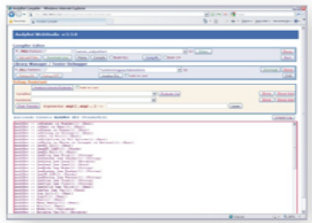
DDL: General Agent Behavior
Manners, Answers, Knowledge

External .NET 2.0 DLL Libraries & Connectors

SMPP SMS interface

e-mail message+files

Robots action/movement



Morphologic Lexer EAGLES 2.0 Tags.

Plain Text

Text Cleansing + Event Filtering-Routing

This cascade makes the 'dirty' job of removing nonsenses, repairing spell errors & detecting input language.

Customer Interface (Text+Files+Actions)

Multiple protocol interface, allowing plain unicode text, peer to peer file transfer, mail notifications, software special events and actions.

DATA XML HTTP SOAP JQUERY

SQL MySQL T-SQL PostgreSQL Oracle

Logging log4net *.txt SysLog UDP

External Events
acoustic sensors
computer vision
accelerometers
user presence
button clicks
mail arrival
file income
+actions

Responses
image animation
sound effects
moving parts
text & voice
responses
mail send
push-file
+actions

App.: AndyBot WebStudio IDE

Alternative Generation

Spellers
PoorMan
Heuristic
Statistic
+
Phonetic
Similarity
Algorithm[2]
+
Smart A.I.
Language
Detection[3]

Morpho+Algorithmic[3] Intelligent Dictionary

Avatar
Virtual Callcenter

Automated & smart callcenters can be built around available commercial text to/from speech interfaces

Hello, I am you Assistant
What Can I do For You?

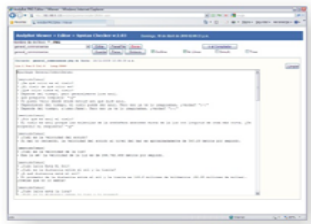
TTS (text to speech)

ASR (speech to text)

MSN chat interface
ICQ chat interface
Gtalk chat interface
AIM chat interface
XMPP chat interface

App.: Live Messenger

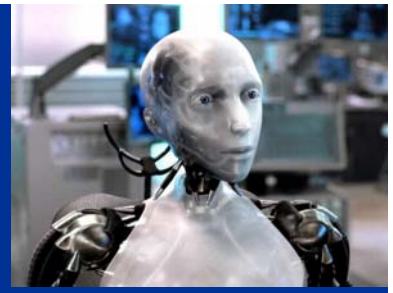
App.: Google Talk



App.: WebStudio DDL Editor

Inteligencia Artificial IV

El Futuro de la PC



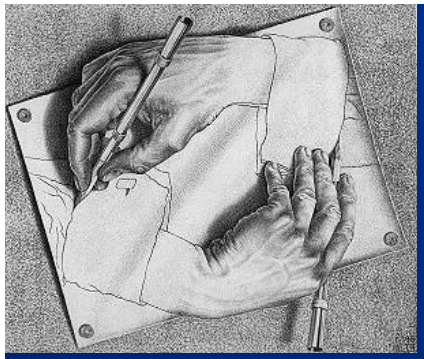
Dentro de 10 años aproximadamente..

- Predicciones (muy probables)
  - Una PC de USD1000 tendrá 100 veces el poder actual.
  - La interfase será más humana (espero y contribuyo)
  - Tendrá más funciones y será de más utilidad que hoy.
  - Tal vez sea todo un conjunto de equipos conectados..

- Deseos (más que predicciones)
  - La futura PC dialogará con inteligentemente con Ud.
  - Será más fácil usarla sin tener que aprender tanto!
    - Que mejore la calidad de vida.
    - Que no sea tan efímera..
    - Que las baterías duren más..

- Delirios (predicciones fantasiosas)
  - La I.A. será mas potente y producirá “La Singularidad”:
    - Superará al intelecto humano (hoy ya lo hace con cuentas..)
    - Logrará autoconocimiento (sabrá de su propia existencia?)
    - Tomará el Control de la Humanidad (aka: Terminator..)

    Y mucho más... predicado por: Kurzweil, Goertzel, etc.



Gracias!

Andrés T. Hohendahl

www.**PandoraBox**.com.ar

andres.hohendahl@fi.uba.ar

web.fi.uba.ar/~ahohenda

Para obtener esta presentación envíe un SMS con la palabra PBOX 1706 al 55588