

# Inteligencia Artificial

## CC52A

Carlos Hurtado L.

[churtado@dcc.uchile.cl](mailto:churtado@dcc.uchile.cl)

<http://www.dcc.uchile.cl/~churtado/>

DCC, Universidad de Chile

# ¿Qué es la Inteligencia Artificial (IA)?

- “The science and engineering of making intelligent machines”. John McCarthy, 1955.
- IA como ciencia (o más bien filosofía)
  - Respuesta a ciertas preguntas
  - **Objeto de estudio**: relación entre capacidades intelectuales y biológicas “superiores” (comportamiento inteligente) y máquinas (computadores).
- IA como tecnología
  - Desarrollo de sistemas
  - Sistemas que realizan tareas avanzadas

# IA como Ciencia: otras definiciones

- “Entendimiento científico de los mecanismos que sustentan el pensamiento y comportamiento inteligente y su implementación en máquinas”.  
*American Association for Artificial Intelligence.*
- “Estudio de facultades mentales mediante el uso de modelos computacionales”.  
*Charniak y McDermott, 1985.*

# IA como Ciencia

- ¿Pueden pensar las máquinas?
  - ¿Qué es una máquina?
  - ¿Qué es la inteligencia?
  - ¿Cómo comprobar que una máquina piensa?
  - ¿Puede un computador pensar?
  - ¿Depende la inteligencia del sustrato físico?

# ¿Cómo comprobar que una máquina piensa?

- Alan Turing, 1950, “Computing Machinery and Intelligence”
- En este artículo, Turing sostiene una idea radical para esos años: en el 2000 existirán computadores inteligentes
- Este enunciado tiene muchas ramificaciones:
  - Ejemplo: ¿Puede existir la conciencia artificial?  
Debate IA débil versus IA fuerte.
- Turing propone un enfoque pragmático: test de “inteligencia” para máquinas.

# Test de Turing

- Participan 3 personas: un hombre (H), una mujer (M) y un interrogador (I)
- Todos están en salas separadas y se comunican mediante un teletipo
- Objetivo de I: preguntar y determinar cuál es el hombre y cuál es la mujer
- Objetivo de H es intentar que I haga una identificación errónea (H gana)

# Test de Turing (cont.)

- En la práctica si H juega muy bien, gana un 70%, si H juega normal gana un 50%.
- ¿Si el reemplazamos H por un computador, este ganará con la misma frecuencia que H?
- Versión moderna: un computador intenta engañar a un interrogador haciéndose pasar por un ser humano.

# Predicción de Turing

*“I believe that in about fifty years' time it will be possible, to programme computers, with a storage capacity of about  $10^9$ , to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning. The original question, "Can machines think?" I believe to be too meaningless to deserve discussion. Nevertheless I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted.”*

# Predicción de Turing: Realidad

- Intentos tempranos: chatbots
  - ELIZA (Weizenbaum 1965).
    - "My mother hates me?"
    - "Why do you say mother hates you".
  - JULIA (Mauldin 1994).
- No se cumplió y está lejos de cumplirse
- Basta ver resultados del torneo *Loebner* 2007
  - <http://www.loebner.net/Prizef/loebner-prize.html>
  - Ningún chatbot ha ganado el *Loebner price* (\$US 100.000)
- Para tener impresión personal, conversar con Joan, chatbot ganador de medalla de bronce torneo 2006 (<http://www.jabberwacky.com/>)

# Resultados Torneo Loebner 2007

**Place - Average Rank - Program - Author(s)**

1st - 3.75 - Joan - Rollo Carpenter

2nd - 3 - Ultra - HalRobert Medeksza

3rd - 2.25 - Cletus - Noah Duncan

4th - 1 - JohnR. - Churchill & M-C Jenkins

# ¿Qué se requieren para aprobar el Test de Turing?

- Procesamiento de lenguaje natural
  - Establecer comunicación satisfactoria en lenguajes humanos
  - Extraer significado e intenciones de lenguaje humano
- Representación del conocimiento
  - Almacenar la información sobre el mundo (modelar el mundo en forma abstracta)
- Aprendizaje
  - Inducción de conocimiento y adaptación a nuevas circunstancias
- Razonamiento
  - Obtención de nuevo conocimiento a partir de conocimiento previo.

# Prueba Total de Turing

- Incorpora interacción física evaluador-computador
  - Procesamiento de imágenes y percepción de objetos
  - Robótica: desplazar objetos

# IA dura vs. IA blanda

- Searle, 1980, IA dura:
  - Inteligencia en el sentido “duro”: estados mentales que alcanzan los entes inteligentes: conciencia, pensamiento y comprensión.
  - Tesis de la IA dura: "the computer is not merely a tool in the study of the mind, rather the appropriately programmed computer really *is* a mind in the sense that computers given the right programs can be literally said to *understand* and have other cognitive states"
- Searle niega la IA dura usando el argumento de la “sala china”.
  - Buen artículo sobre esto: the internet Encyclopedia of Philosophy: <http://www.iep.utm.edu/c/chineser.htm>
- Los computadores son sólo manipuladores de símbolos que en el mejor de los casos sólo simularán la mente humana.

# IA dura vs. IA blanda (cont.)

- Searle, 1980:
  - La materia de que estamos hechos es fundamental para la Inteligencia.
  - El pensamiento sólo puede ocurrir en máquinas vivientes hechas de proteínas.
- Newell y Simon, 1976:
  - Hipótesis del sistema físico de símbolos: máquina que es capaz de manipular datos simbólicos bastaría para emular inteligencia.
  - La inteligencia es independiente del sustrato físico

# AI como Tecnología

- “Tarea de lograr que los computadores piensen... *máquinas con mente*, en su amplio sentido literal.” Haugeland, 1985.
- “El arte de crear máquinas con capacidad de realizar funciones que ejecutadas por personas requieren inteligencia.” Kurzweill, 1990.
- Técnicas para la resolución de problemas complejos.
- Enfoque de este curso: métodos para desarrollar programas avanzados.

# Modelos para la IA

- Basados en procesamiento de símbolos
  - Hipótesis de sistema físico de símbolos de Newell y Simon, 1976.
  - Solucionador General de Problemas
  - Sistemas Expertos: aplicación de operaciones lógicas sobre bases de conocimiento (sentencias declarativas)
- Sub-simbólico:
  - Concepto de señal en lugar de símbolo
  - Inteligencia se desarrollo en muchos años de evolución
  - Primero debemos concentrarnos en emular las capacidades de procesamiento de señal que poseen los animales más simples
  - Ejemplo: Redes Neuronales

# Pre-Historia de la IA

- Aristóteles: estructuras de argumentación lógica (e.g., silogismos)
- 1642. Pascal: Pascalina, calculadora mecánica de 8 dígitos
- 1642. Leibnitz. Calculador y diseño de primer "razonador"
- 1832. Babbage: Analytical Engine, computador programable mecánico.
- 1847. George Boole: lenguaje formal para inferencia lógica, base de la lógica de proposiciones
- 1879. Frege: formalización de la lógica de primer orden
- 1930. Godel: completitud de axiomas de lógica de primer orden

# Historia de la IA

- 1921. Karel Capek. Término “robot”
- 1928. Von Neumann. Teorema y algoritmo Minimax
- 1937 Turing. Turing Machine.
- 1938. Turing, Church: Tesis de Church.
- 1945. Von Newman. Diseño de la arquitectura básica de los computadores actuales
- 1945 Mauchley y Eckert. ENIAC (Electronic Numerical Integrator and Calculator) Primer computador programable 100% electrónico.
- 1950. Turing: Test de Turing para reconocer inteligencia artificial
- 1950. Minsky y Edmonds: Primeras redes neuronales
- 1951. EDVAC. Primer computador de Von Newman
- 1955. Newell, Simon y Shaw. Logic Theorist. Primer demostrador automático de teoremas

# Historia de la IA (cont.)

- 1955. McCarthy. Acuña el término Inteligencia Artificial
- 1956. IBM 701. Primer computador electrónico en el mercado.
- 1957 Newell y Simon Solucionador General de Problemas:
  - integración simbólica, problemas de álgebra, control de robots.
- 1957. Samuel. Primer jugador artificial de ajedrez competitivo.
- 1958. McCarthy: Lenguaje de programación LISP
- 1958 McCarthy: Sistema Consejero: sistema declarativo genérico basado en cálculo de predicados para la resolución de problemas.
- 1959 Gelernter. Demostrador de Teoremas de Geometría de IBM
- 1965. Feigenbaum y Lindsay. DENDRAL, primer sistema experto.
  - Predicción de la estructura de moléculas orgánicas a través de su fórmula química. Se hizo inmanejable al poco tiempo.
- 1965. Weizenbaum. ELIZA.

# Historia de la IA (cont.)

- 1965-1970. Minsky y Papert. Proyecto Blocks Microworld en MIT.
- 1965. Simon: Los computadores serán capaces de realizar “any work a man can do” en 1985.
- 1966. National Research council de USA suprime todo financiamiento a la investigación en traducción automática de lenguaje humano.
- 1968. Kubrick y Clarke. 2001: A Space Oddyssey. Computador que enloquece (HAL).
- 1971. Colmerauer y Rousel. PROLOG.
  - Revisado en 1974 para operar con cláusula de Horn.
- 1974. Shortliffe y Buchanan. MYCIN (Buchanan): sistema experto para diagnóstico de infecciones del sistema linfático.
- 1978 Winograd. SHRDLU: comprensión de lenguaje natural

# Historia de la IA (cont.)

- 1979. McDermott Primer sistema experto comercial. XCON.
- 1980-1985 Primera generación de sistemas expertos en el mercado.
  - XCON (DEC, ayuda para configurar sistemas VAX), pasó de 300 a más de 3000 reglas.
  - PROSPECTOR (prospección de minerales),
  - SE de BELL para análisis de errores en redes
  - FOLIO (asesor de inversiones)
  - WILLIARD (predictor de tormentas)
- 1986 se concretan ventas de \$US 425 millones en software (LISP, PROLOG) y hardware especializado IA. Surgen diversas start-up de IA.
- 1985. Primeros sistemas comerciales de visión y reconocimiento de lenguaje hablado.
- 1985. Surgen sistemas aplicados basados en redes neuronales.

# Historia de la IA (cont.)

- 1987 en adelante. Sistemas expertos comienzan a mostrar falencia en su uso comercial:
  - Tamaño inmanejable, reglas superpuestas (XCON alcanza 10.000 reglas)
  - Inflexibilidad: incapacidad de aprender
  - Costo de actualización (costo de elicitar conocimiento)
  - Resulta más barato entrenar humanos que manejar sistemas expertos.
  - Firebaugh: “expert systems are limited to any problem that can be and frequently is solved by your in-house expert in a 10 to 30 minute phone call,”

# Historia de la IA (cont.)

- 1990-2000. Desarrollo de aplicaciones prácticas de IA:
  - Software de programación
  - Reconocimiento de lenguaje
  - Reconocimiento de caras
  - Sistemas de clasificación de riesgo financiero
  - Sistema de detección de fraude
  - Minería de Datos
  - etc....

# Historia de la IA (cont.)

- 1995. SKYCAT: Descubrimiento de 16 cuasares usando minería de datos.
- 1996. Kasparov 4 - IBM Deep Blue 5
- 1997. Kasparov 2.5 - IBM Deep Blue 3.5
  - Red de 200 cpus que analiza 200 millones de movidas por segundo (vs. 3 de Kasparov)
- 1997. Más de 50 equipos participan en el torneo de fútbol RoboCup.
- 1999. Sony introduce robot mascota AIBO.
- 2002. MIT AI Lab. iRobot: robot-aspiradora
- 2005. *Volkswagen Touareg R5* gana el “DARPA Grand Challenge” al completar conducción automática en 7 horas de un trayecto de 121 km. Sólo 3/27 vehículos llegaron a la meta.
- 2006. Venta de 2 millones de iRobots.

# Historia de la IA

- “En la actualidad es imposible conversar con un computador en un cocktail. Sin embargo, en alguna manera, la IA ya está con nosotros: en el control de crusero de autos, los servidores que rutean email, y en los avisos personalizados que estorban la ventana de nuestro navegador Web.”  
– Jennifer Kahn. *Wired* (March 2002/10.03).

# Programa del Curso

- Contenido
- Evaluación
- Referencias

# Contenido del Curso

## 1. Representación de Conocimiento

- Lógica de proposiciones
- Lógica de primer orden
- PROLOG y sistemas expertos
- Representación de conocimiento en la Web
  - Redes semánticas (RDF, RDFS)
  - Ontologías (OWL).

## 2. Solución de problemas mediante búsqueda

- Modelación de problemas como búsqueda
- Algoritmos de búsqueda
- Juegos

# Contenido del Curso (cont.)

## 1. Aprendizaje de Máquinas

- Sistemas de reconocimiento de patrones
- Aprendizaje supervisado
  - Árboles de decisión, redes Bayesianas, redes neuronales, discriminantes lineales.
- Aprendizaje no supervisado
  - Clustering

# Evaluación

- Dos controles y un examen final
- Dos controles de lectura y una nota de participación en clases
- Tres tareas prácticas
- Nota final:
  - 60% nota controles y examen.
  - 40% nota de tareas, controles de lectura y participación en clases.

# Referencias

- Apuntes del curso
- Libros:
  - Inteligencia Artificial: un enfoque moderno. S. Russell, P. Norvig. Prentice Hall, 1995. (En Biblioteca)
  - Inteligencia Artificial: una nueva síntesis. N. Nilsson. Mc Graw Hill, 2001.
  - Machine Learning. T. Mitchell. McGraw Hill, 1997.
  - Pattern Classification. R. Duda, P. Hart, D. Stork. Wiley Interscience

# Control de Lectura 1

- Jueves 29 de Marzo, horario de clases.
- ¿What is Artificial Intelligence? John McCarthy
  - <http://www.formal.stanford.edu/jmc/whatisai/whatisai.html>
- Alan Turing. Computing Machinery and Intelligence. Mind, 1950.
  - <http://www.loebner.net/Prizef/TuringArticle.html>
- Argumento y contrargumentos de la sala china de John Searle.
  - Fuente: Web.