

# 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 general:**
  - Respuesta a ciertas preguntas
  - Relación entre capacidades intelectuales y biológicas “superiores” (comportamiento inteligente) y máquinas (computadores).
- **IA como tecnología (este curso):**
  - Desarrollo de sistemas computacionales que realizan tareas avanzadas
  - Modelos computacionales, algoritmos, complejidad de problemas.

# Otras Definiciones de IA

- “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?

# IA como Disciplina Computacional

- Aprendizaje de Máquinas
- Lógica y Representación del Conocimiento
- Procesamiento de Lenguaje Natural
- Procesamiento de señales y robótica
- Otras disciplinas: complejidad, grafos, optimización.

# ¿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 dependiendo de la definición de inteligencia y de máquina.
  - Ejemplo: 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
- I gana si al cabo de un número fijo de preguntas identifica correctamente a la mujer
- En caso contrario, gana H.

# Test de Turing (cont.)

- Ejemplo:
  - I: Will H please tell me the length of his or her hair?
  - H: My hair is shingled, and the longest strands, are about nine inches long.
- ¿Si el reemplazamos H por un computador, este ganará con la misma frecuencia que H?

# ¿Cuándo un Computador Pasa el Test de Turing?

- En la práctica si H juega muy bien, gana un 70%, si H juega normal gana un 50%.
- Un H que muestra un grado normal de “inteligencia” gana un 30% de la veces a interrogadores con un grado normal de “inteligencia”.
- Es decir, un computador para pasar el test de Turing debe engañar más de un 30% de las veces al interrogador.

# Aprendizaje

- Según Turing, el interrogador (I) puede ser una persona de cualquier sexo.
- Si es I mujer, será mucho más fácil para I ganar, formulando preguntas que requieran conocimiento específico que sólo tengan las mujeres.
  - Si no permitimos que H (o el computador) aprenda sobre mujeres previamente, estaríamos midiendo conocimiento en lugar de inteligencia.
- Si I es un hombre, entonces no sirve basar sus preguntas en conocimiento específico de las mujeres.
  - En este caso no tenemos el problema anterior.
- La primera crítica al test de Turing es que hay un sesgo por conocimiento y no inteligencia y no mide capacidad de aprendizaje.

# Versión Popular del Test de Turing

- Un interrogador, un computador y un humano.
- El interrogador debe identificar al humano.
- Se argumenta que esta versión desfavorece al computador
- I puede intentar medir conocimiento o experiencia en lugar de inteligencia:
  - ¿Mujer que tuvo romance con Bill Clinton?

# Test Total de Turing

- Incorpora interacción física evaluador-computador
- Procesamiento de imágenes, percepción de objetos
- Robótica: por ejemplo, desplazar objetos.

# ¿Qué se requiere para pasar el Test de Turing?

- **I:** In the first line of your sonnet which reads 'Shall I compare thee to a summer's day', would not 'a spring day' do as well or better?
- **H/C:** It wouldn't scan.
- **I:** How about 'a winter's day' That would scan all right.
- **H/C:** Yes, but nobody wants to be compared to a winter's day.

# ¿Qué se requiere para pasar el Test de Turing?

- **I:** Would you say Mr. Pickwick reminded you of Christmas?
- **H/C:** In a way.
- **I:** Yet Christmas is a winter's day, and I do not think Mr. Pickwick would mind the comparison.
- **H/C:** I don't think you're serious. By a winter's day one means a typical winter's day, rather than a special one like Christmas.

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

- **Procesamiento de lenguaje natural**
  - Establecer comunicación satisfactoria en lenguajes humanos
  - Comprender significado e intenciones del lenguaje humano
- **Representación de conocimiento y lógica**
  - Representar en forma abstracta información sobre el mundo
  - Obtener nuevo conocimiento a partir de conocimiento previo.
- **Aprendizaje**
  - Inducción de conocimiento de la experiencia
  - Adaptación a nuevas circunstancias

# Profecía de Turing

- En el año 2000 será posible programar computadores con capacidad de memoria de  $10^9$  que ganarán más de un 30% de las veces al interrogador.
- En esos años, la gente educada hablará de la existencia de máquinas inteligentes sin esperar que alguien los contradiga.

# Predicción de Turing: Realidad

- No se cumplió y está lejos de cumplirse
- Intentos tempranos: chatbots
  - ELIZA (Weizenbaum 1965).
    - "My mother hates me?"
    - "Why do you say mother hates you".
  - JULIA (Mauldin 1994).

# Predicción de Turing: Realidad (cont.)

- Resultados del torneo *Loebner 2007*
  - Ningún chatbot ha ganado el *Loebner price* (*\$US 100.000*)
  - <http://www.loebner.net/Prizetf/loebner-prize.html>
- Joan: chatbot ganador de medalla de bronce 2006 (<http://www.jabberwacky.com/>)

# Resutados 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



# Chatbots

- Sin embargo, si la gente no está avisada que puede estar conversando con un chatbot, en general cree que conversa con un humano.
- Según muchos, esto prueba que los chatbots ya aprueban el test de Turing original (“Female Turing Test”).

# 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: los computadores son manipuladores de símbolos que en el mejor de los casos sólo simularán la mente humana.

# Argumento de la Sala China

- Searle pone en duda la capacidad del test de Turing de medir “inteligencia” usando el argumento de la “sala china”.
  - Referencia: the internet Encyclopedia of Philosophy:  
<http://www.iep.utm.edu/c/chineser.htm>

# Argumento de la Sala China

- Un hombre en una sala
- Un manual muy completo con instrucciones sobre cómo responder preguntas en Chino.
- Una persona que no sabe nada de Chino en esta sala.
- Esta persona puede engañar a cualquier Chino que sabe Chino.

# 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

# Kasparov vs. Deep Blue

- IBM desarrolló durante 1 año a Deep Blue en completo secreto.
- Implementado sobre un RS/6000 SP, con capacidad de procesamiento equivalente a 256 PCs medianos. Escrito en C sobre AIX
- En el año 2006 era el 259avo computador más potente en el mundo.
- Uso de aprendizaje de máquinas para afinar la función de utilidad.
- En 1996 Kasparov vence a Deep Blue 4-2
- Esta versión procesaba ~100 millones de jugadas por segundo
- Su gran deficiencia era el juego de apertura

# Kasparov vs. Deep Blue

- En 1997 se enfrentan nuevamente
- Deep Blue mejora a ~200 millones de jugadas por segundo
- Implementa una librería de 1 millón de apertura (incluidas todas las jugadas de Kasparov) analizadas por tres grandes maestros
- Resultado: Deep Blue 3 Kasparov 2
- Kasparov comenta: "...por fin la cantidad se transformó en calidad...".
- Kasparov ha llegado a decir que había una mente humana detrás de Deep Blue y reclama que no pudo estudiar partidas de Deep Blue antes del juego.
- Reclama que Deep Blue era mejorado entre partidas y pide los logs del servidor que no son entregados por IBM.

# Kasparov vs. Deep Blue

- Gran éxito publicitario para IBM (\$US 100 millones aprox gratis en cobertura periodística).
- Kasparov pide otro enfrentamiento pero IBM se niega y retira a Deep Blue.

# IA como Tecnología

- “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).

# 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.

# 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 (predicador 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
- 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.

# 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
  - Segmentación

# 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