Robots and Language: A Practical Perspective

D. Butterworth[1]

Robots are moving out of the factory and beginning to enter our daily lives. Autonomous robots have demonstrated advanced functionality in structured environments, but they fail in ambiguous, un-structured environments like a house. To enable robots to work amongst us, and operate safely in our homes, they need to be able to communicate effectively with humans. Researchers are investigating robot language systems based on human language abilities, using a complex, integrated system of symbol representations grounded in a cognitive map. By debating the problems with taking this human-inspired approach towards language in robots, this paper presents the alternative perspective that a robot's language capabilities need only be sufficient enough that it can be operated in a natural and straightforward manner. Ultimately, a robot with practical language abilities will be able to comprehend natural, spoken commands from a human, and provide verbal feedback to its operator.

Introduction

Robotics is widely used in manufacturing industries, but practical robots have not yet found their way into our daily lives. Science fiction has sketched a future filled with advanced technology and by the twenty-first century it was believed that robots would already be common-place in our daily lives (Levitan & Johnson 1982). With advancements in DC motor technology, battery storage capacity and computer processing power, the consumer and domestic robotics industries *are* expanding (Global Industry Analysts Inc. 2010) but there are some fundamental issues that need to be addressed before robots can work effectively together with humans and operate safely in our homes (Weng, Chen & Su 2009). Robots will need advanced language capabilities so that humans can communicate with them in a natural manner. This paper will debate the problems with taking a **cognitive robotics** approach towards language development and present the alternative perspective that a robot's language abilities need only be sufficient enough that it can be operated in a straightforward and natural manner.

Robotics today

Most robots currently in use are either movement-repeating machines or tele-operated, remote-control devices. Industrial robots have advanced vision capabilities (Espingardeiro 2010), can operate at high speeds (Eaton 2008), and handle very heavy (Diaz 2009) or microscopic objects (Gu et al. 2009). Tele-operated robots are controlled directly by a human operator and include medical robots (Leven et al. 2005) and maintenance inspection robots (Guizzo 2010). These types of robots are used in static, **structured environments** and require only a low-level of computational intelligence; if the environment doesn't change then the robot can operate blindly without problems.

More advanced robots are fully **autonomous machines** that can operate in dynamic, unknown, and **unstructured environments**. This means that a mobile robot can be programmed with a specific task, placed in an unfamiliar location, successfully navigate through a changing environment with unknown obstacles, and finally achieve its goal. Military research has created robots that can negotiate uneven terrain (Raibert et al. 2008) and navigate at high-speeds in an unknown environment (Thrun et al. 2006). Other autonomous mobile robots can swim underwater (Hirata 2000) or balance on two wheels (Nguyen et al. 2004). But a key problem is, that while we have built machines that can negotiate almost any terrain on earth, we know much less about creating the **artificial intelligence** that is *the* core of any advanced robot, and one of the most challenging environments this intelligence can be tested is in the unstructured environment of the home.

---

[1] Undergraduate student, University of Queensland, Brisbane, Australia

Robotics in the home

Domestic robots have been around since the 1980's, but due to technological limitations they have functioned primarily as entertainment or educational devices (Elmaghraby & Jagannathan 1985) and toys (ShanieAIBO). More recently, a range of companies have released robotic vacuum products (Forlizzi & DiSalvo 2006), and other more expensive robots have been designed especially for the household, utilising a humanoid platform (Sofge 2008) or tracked-wheel design (Glaskowsky 2007). But all of these robots still have a relatively low-level of intelligence (FIG. 1).

Currently the domestic robotics market is very small and primarily focused on robot toys, but the market will expand as countries seek ways to alleviate their aged-care problems (Dethlefs & Martin 2006) and also due to consumer demand for time-saving devices. In the push to create intelligent, autonomous household robots, there are issues that need to be addressed regarding robots working in close proximity to humans, and questions about human interaction with these artificially intelligent machines, but this paper is focused on the need to develop language capabilities for robots.

Language and robots

Robots will need advanced language capabilities so that humans can operate them using natural, verbal commands, such as "*robot, could you help me with the groceries*". The robot would also be able to talk and provide feedback to the user, such as *"I think you want me to lift these bags out of the car and bring them inside the house, is that correct?"* The importance of natural language instructions, rather than specific single-word directives like *"robot move"*, is that it makes the robot easy to use without the operator modifying their own behaviour.

This type of robot would be an autonomous mobile robot, with some sort of artificial intelligence, capable of performing advanced goal-oriented behaviour inside the home. But if a robot is essentially a computer attached to a mechanical platform, it raises the question as to why language should be studied with robots, as opposed to developing language capabilities through a stand-alone computer software system. This is where the field of cognitive robotics differs from a computer science approach, suggesting that language is one part of an integrated system and therefore robots can't develop language abilities without also having the sensory attachments that allow them to see, touch and experience the world they inhabit.

The study of language and robots is a cross-disciplinary field, drawing from knowledge in areas such as;

Engineering,
to build an electro-mechanical robotic platform;

Computer Science,
to create an artificially intelligent computer brain, Speech Recognition software, to comprehend spoken words, and Speech Synthesis, so the robot can talk;

Cognitive Science,
taking our understanding of the human mind and intelligence, and using that to build a robotic brain;

Figure 1 | Domestic robots: Mahru-Z (K.I.S.T.) Korea and PR2 (Willow Garage) U.S.
The most advanced domestic robots are highly complex mechanical devices, with significant computational power, but still lack anything that could be described as "intelligence".

COGNITIVE ROBOTICS
Building robots with intelligence based on research into the human mind and thought, as opposed to traditional Artificial Intelligence methods.

STRUCTURED ENVIRONMENT
Where all environmental attributes are controlled or generally predictable. e.g. Industrial robots operate in restricted "cells" with discrete lighting conditions.

UN-STRUCTURED ENVIRONMENT
Containing unpredictable variables, including varying daylight, unknown obstacles, and humans.
e.g. the Home.

AUTONOMOUS ROBOT
A robot that can perform goal-orientated tasks in an un-structured environment, autonomously with no human intervention.

ARTIFICIAL INTELLIGENCE
The field of Computer Science that aims to create the appearance of intelligence in machines.

and, *Linguistics*,
which shows us that language is more than just words and is integrated with all human cognitive processes, and to generalise new research findings across other non-English languages.

Literature review

The rudimentary approach to adding language capabilities to a robot is using speech recognition technology. The robot is trained to recognise an operator's voice and matches known voice commands with actions, using a look-up table. Aside from the problems that still exist with Speech Recognition technology, the key problem with this type of system is that it requires large, rigid data structures that don't correspond with the fluidic way that humans use their language syntax. Roy, Kai-Yuh & Mavridis (2004) programmed a robot arm manipulator to respond to spatial and action words, for example "*lift, block, behind*". Skubic et al. (2004) programmed a mobile robot to understand spatial words and make statements about its environment, such as "*the object is in front of me*".

Other research uses a human-inspired, cognitive science approach, which takes our knowledge of how concepts are represented in a human brain and applies them to build more advanced robotic brains. The robot is taught some words, and attaches these, along with visual input from its camera, to a **neural network**; a sort of **cognitive map** for a robot. Steels (2001) proposed the idea of playing structured games with robots, in order to teach them language, and has demonstrated this using a range of platforms like the Sony AIBO robot dog and Sony SDR Humanoid robot. It has also been shown that robots can be taught a small set of basic action words and combine them to develop new combination words, in the experiment by Cangelosi et al. (2007) with robot arms, and with the iCub humanoid robot (Cangelosi 2010).

The aforementioned research has been based around a human teaching language to a robot, but cognitive science and language evolution have also been combined to examine how language develops between multiple robots, without any human input at all. This extends the theory that complex language abilities must be part of a robot's integrated cognitive architecture and proposes that robots must learn language on their own, so the language that evolves is not tainted by any meaning implied by a human teacher. Baillie and Nottale (2005) have shown how a pair of Sony AIBO robot dogs can play games to evolve their own language. Similarly Steels (2008) has presented Sony Qrio humanoid robots that played a colour guessing game and evolved their own words for different colours. In one of the only studies using autonomous mobile robots, Schulz et al. (2010) have demonstrated robots that can explore an indoor area, create a language of place names to describe the environment and evolve new words to describe distance, direction and time, along with new, imagined place names outside the robot's immediate environment. (FIG. 2).

The human-inspired approach

After the lack of success with simple word recognition systems, where each word is directly attached to a singular

---

NEURAL NETWORK
A computer program that mimics the behaviour of a network of biological neurons in the human brain.

COGNITIVE MAP
A programming construct inspired by cognitive science, that allows a robot to process information about its environment, recall memories and solve problems.

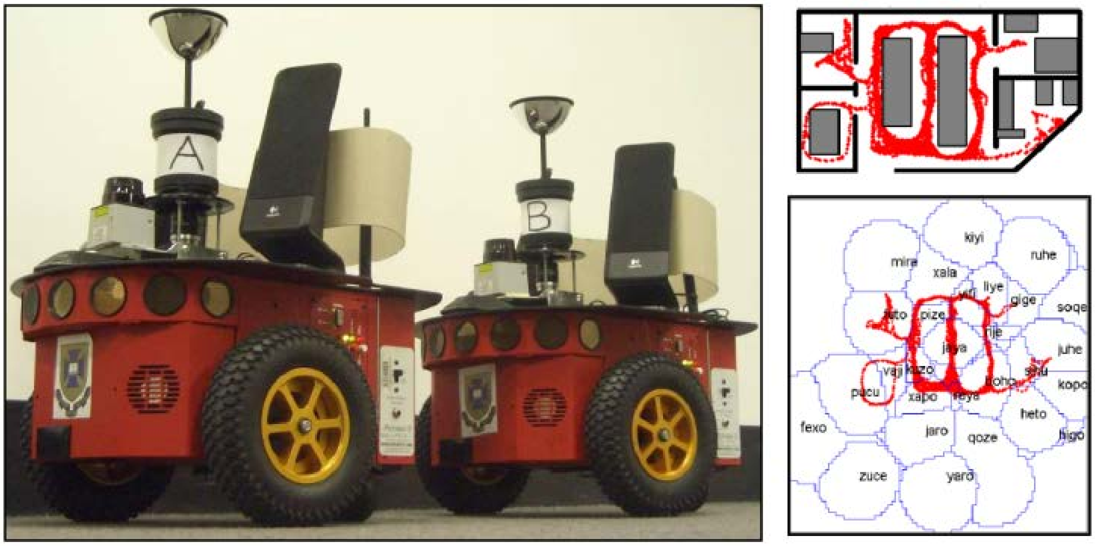

Figure 2 | The evolutionary approach to language learning in robots (Schulz et al. 2010)
These robots can create their own language, but practical household robots must be capable of understanding human language from the moment they are first used.

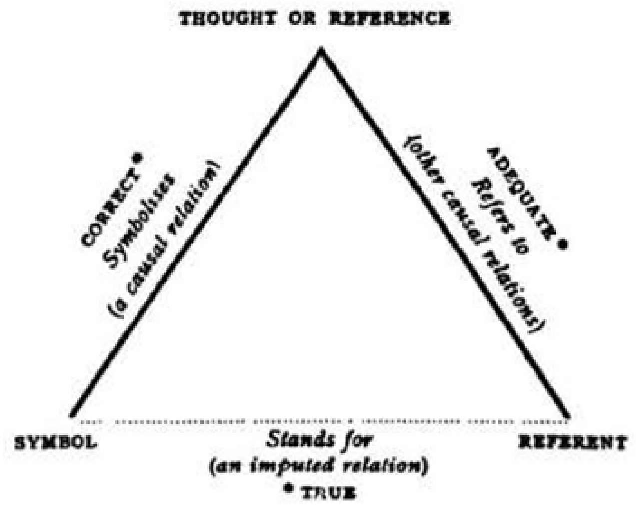

Figure 3 | The Triangle of Meaning (Ogden & Richards 1923)
Also called the "Semiotic Triangle", describes the symbol grounding process. An object or referent in the real world is translated into a conceptual symbol in the human mind, and further translated into thoughts and meaning that are associated with the object.

meaning, researchers are now taking a human-inspired approach. This means trying to recreate a human's language abilities in a robotic brain, which would require solving the deep philosophical and cognitive science problems of **symbol grounding** (Harnad 1990) and the Robotic Turing Test **(**Harnad & Scherzer 2008)**.**

The Symbol Grounding Problem is a question in cognitive science relating to how symbols in the real world are related to meaning inside the mind. This problem has been extensively discussed by Harnad (1990) and originally presented by Ogden and Richards (1923) as the "Triangle of Meaning" (FIG. 3), the philosophy of how symbols relate to objects. If a human and a robot are looking at a red ball, then the actual physical ball is called the object or referent. The name of the object, "a red ball", is the symbol; the human will have learnt this as a child, and the robot will have been programmed to call it a "red ball". The key issues with implementing this in a robot, however, are that;
the human and the robot will each have an individual, internal representation of the ball, called the "thought", and;
both the "thought" and the "symbol" are both attached back to the original object using a system that allows translation between the object and the internal representation, and this ability is called grounding.

Successfully implementing this in a robot would, it is believed, solve the Robotic Turing Test (Harnad & Scherzer 2008). This is where a human could interact with a robot and not be able to distinguish it from another human, excluding obvious visual give-aways. This differs from the general Turing Test (Turing 1950), where a human converses with a computer through a PC terminal, and is an important difference, as the ability of a robot to "ground" means that it can take a word and translate that back into its sensory representation of whatever object is being referred to in the real world. So the problem of giving language ability to robots is now considered to be inextricably linked to these deeper problems regarding how a robot could translate the particular context of a word, determine the appropriate meaning of each word, and relate the appropriate **affordances** or actions that are usually associated with certain words. Having started with crude word recognition, it seems the answer to language in robots is to replicate a complex, integrated cognitive architecture inside a robot, and that's certainly no easy task.

SYMBOL GROUNDING
The cognitive process that connects real world objects to thoughts and meaning in the brain.

AFFORDANCE
A theory in psychology and perception studies, suggesting that each object has a particular set of action possibilities. e.g. that a ball has the "affordance" of being bounced.

A practical perspective

Inspired by mythology and science fiction, humans have for a long time tried to create a robotic man or android (Childress 2000), but such a creation has so far exceeded our technical know-how. Successful humanoid robots have only been realised relatively recently in robotics history and, despite their literary inspiration, they were logically regarded as *the* most effective design for a robot that needs to operate in a human environment, containing stairs, rough surfaces and uneven ground. But in parallel to these decades of humanoid robot research, a myriad of alternative locomotion solutions have been investigated and now, due to advances in engineering, the humanoid robotic platform is no longer uniformly regarded as the best solution to a robot in the home. So if we're no longer trying to build a robotic man, then in trying to build robots that can effectively communicate with humans, why are we trying to replicate human language abilities?

The debate continues regarding how primitive humans first gained language abilities beyond their primate ancestors. But this is irrelevant for developing practical robots; the goal is not to create a new robotic race that learns and evolves its own unique robotic language dialect. The goal is to create robots that can comprehend common human languages with sufficient ability to complete goal-orientated tasks. A practical robot should be capable of learning language and other tasks, in a similar fashion to how a person learns a new hobby; that there is some new knowledge to be acquired, but the fundamentals are pre-programmed, which is in stark contrast to a biologically-inspired robot that must learn through evolution, like a caveman trying to make fire.

But robotics has often taken inspiration from nature; after all, science says that each of the diverse range of life-forms are nothing more than complex biological systems, so if we can learn enough about how they work then it would be comparatively trivial to replicate such a system with circuits and wires. Roy (2009) has taken this approach and recently completed a ground-breaking project to record a comprehensive audiovisual record of the first three years of a child's life, focusing on how the child acquires language. The goal is to feed this data into a robot, trialling a range of learning algorithms, to examine what the robot is able to learn. But even if we do succeed in creating a cognitive architecture advanced-enough to give complex language abilities to robots, is this ultimately a desirable outcome? Perhaps we are asking the wrong questions in the quest to enable language in robots.

The successful outcome of building an advanced, cognitive framework that allows a robot to acquire language is that it would be capable of performing the grounding process, described earlier in this paper. For example, given an instruction to "*go and clean the kitchen*" the robot would successfully translate those word symbols into an array of internal representations, thereby associating all the possible meanings of the verbal command. But the logical outcome of such an ambiguous instruction is an equally ambiguous or confusing response: "*where must the robot go*", "*how is the kitchen to be cleaned*", and "*which kitchen is being referred too*". What is considered to be natural human language is fundamentally far too ambiguous to be useful for a robotic machine; this is demonstrated by a common situation in which three people are given a verbal instruction and all three derive a different interpretation of that instruction, yet each interpretation is equally logical and accurate. Such ambiguity would be a hindrance to a practical robot, and while verbal commands should appear natural to the human operator, the robot should interpret them using a rigid system with defined, predictable outcomes.

Conclusion

For robots to be able to operate safely in our homes, and in close-proximity to humans, they will need the ability to communicate using language. To enable the average user to operate a robot, with minimal training, this language must be natural and easy to use. By debating the problems with building a robot language system based on complex, human language abilities, this paper has proposed an alternative practical perspective, that a robot's language abilities need only be sufficient enough for it to comprehend straight-forward verbal commands. Robots are tools and there is little purpose in attempting to build mind-reading machines that can compensate for ambiguous instructions given by a human operator. By asking the correct design questions, practical household robots with both advanced functionality *and* human-friendly interfaces will be realised.

References

Baillie, J-C & Nottale, M 2005, 'Talking Robots: a Fully Autonomous Implementation of the Talking Heads', in *IJCAI-05, Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence*, Edinburgh, Scotland, UK, pp. 1725-6.

Cangelosi, A, Tikhanoff, V, Fontanari, JF & Hourdakis, E 2007, 'Integrating Language and Cognition: A Cognitive Robotics Approach', *IEEE Computational Intelligence Magazine*, vol. 2, no. 3, pp. 65-70.

Cangelosi, A 2010, 'Grounding language in action and perception: From cognitive agents to humanoid robots', *Physics of Life Reviews*, vol. 7, no. 2, pp. 139-51.

Childress, DH 2000, *Technology of the Gods: The Incredible Sciences of the Ancients*, Adventures Unlimited Press, Illinois.

Dethlefs, N & Martin, B 2006, 'Japanese technology policy for aged care', *Science and Public Policy*, vol. 33, no. 1, pp. 47-57.

Diaz, J 2009, *World's Most Powerful Industrial Robot Can Lift Entire Car Frames*, 26 February, Gizmodo, viewed 23 October 2010, < http://gizmodo.com/5161284/worlds-most-powerful-industrial-robot-can-lift-entire-car-frames-tear-apart-several-humans-at-once>.

Eaton, K 2008, *ABB FlexPicker Robot's Legs Move So Fast It's Scary*, 25 March, Gizmodo, viewed 23 October 2010, <http://gizmodo.com/371917/abb-flexpicker-robots-legs-move-so-fast-its-scary>.

Elmaghraby, AS & Jagannathan, V 1985, 'Study of the Hero-1 educational robot', *Transactions of CoED Journal*, vol. V, no. 3, pp. 20-24.

Espingardeiro, A 2010, *Will Robots Pick Your Grapes One Day?*, 1 February, IEEE Spectrum, viewed 23 October 2010, <http://spectrum.ieee.org/automaton/robotics/industrial-robots/will-robots-pick-your-grapes-one-day>.

Forlizzi, J & DiSalvo, C 2006, 'Service Robots in the Domestic Environment: A Study of the Roomba Vacuum in the Home', *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*, Salt Lake City, United States, pp. 258-65.

Glaskowsky, P 2007, *Anybots: Building your robot army*, 27 September, CNET, 23 October 2010, <http://news.cnet.com/8301-13512_3-9785928-23.html>.

Global Industry Analysts Inc. 2010, *Global Service Robotics Market to Reach US$38.42 Billion by 2015 According to New Report by Global Industry Analysts*, press release, July 12, Global Industry Analysts Inc., viewed 23 October 2010, <http://www.prweb.com/releases/service_robotics/professional_robots/prweb4240924.htm>.

Gu, H, Chao, J, Xiao, SJ & Seeman, NC 2009, 'Dynamic Patterning Programmed by DNA Tiles Captured on a DNA Origami Substrate', *Nature Nanotech*, vol. 4, no. 4, pp. 245-9.

Guizzo, E 2010, *Japanese Snake Robot Goes Where Humans Can't*, 17 October, IEEE Spectrum, viewed 23 October 2010, <http://spectrum.ieee.org/automaton/robotics/industrial-robots/japanese-snake-robot-goes-where-humans-cant>.

Harnad, S 1990, 'The symbol grounding problem', *Physica D: Nonlinear Phenomena*, vol. 42, pp. 335-46.

Harnad, S & Scherzer, P 2008, ‘First, Scale Up to the Robotic Turing Test, Then Worry About Feeling’, *Artificial Intelligence in Medicine*, vol. 44, no. 2, pp. 83-89.

Hirata, K 2000, ‘Development of Experimental Fish Robot’, *Proceedings of the Sixth International Symposium on Marine Engineering (ISME)*, Tokyo, Japan, pp. 711-714.

K.I.S.T., *Korea Institute of Science & Technology (K.I.S.T.)*, viewed 23 October 2010, <http://www.kist.re.kr/en/index.jsp>.

Leven, J, Burschka, D, Kumar, R, Zhang, G, Blumenkranz, S, Dai, X, Awad, M, Hager, GD, Marohn, M & Choti, M 2005, ‘DaVinci Canvas: A Telerobotic Surgical System with Integrated, Robot-Assisted, Laparoscopic Ultrasound Capability’, *MICCAI 2005 - Medical Image Computing and Computer-Assisted Intervention - Lecture Notes in Computer Science*, vol. 3749, pp. 811-8.

Levitan, SA & Johnson, CM 1982, ‘The future of work: does it belong to us or to the robots’, *Monthly Labor Review*, vol. 105, no. 9, pp. 10-14.

Nguyen, HG, Morrell, J, Mullens, K, Burmeister, A, Miles, S, Farrington, N, Thomas, K & Gage, D 2004, ‘Segway robotic mobility platform’, *Proceedings of Mobile Robots XVII*, Philadelphia, United States, pp. 207-20.

Ogden, CK & Richards, IA 1923, *The Meaning of Meaning*, 8th edn, Harcourt, Brace & World, Inc., New York.

Raibert, M, Blankespoor, K, Nelson, G & Playter, R 2008, ‘BigDog, the Rough-Terrain Quadruped Robot’, *Proceedings of the 17th World Congress, The International Federation of Automatic Control*, Seoul, Korea, pp. 10822-5.

Roy, D, Kai-Yuh, H & Mavridis, N 2004, 'Mental imagery for a conversational robot', *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 34, no. 3, pp. 1374-83.

Roy, D 2009, ‘New Horizons in the Study of Child Language Acquisition’, *Proceedings of Interspeech 2009: 10th Annual Conference of the International Speech Communication Association*, Brighton, UK, pp. 13-20.

Schulz, R, Glover, A, Wyeth, G & Wiles, J 2010, ‘Robots, Communication, and Language: An Overview of the Lingodroid Project’, paper to be presented at the Australasian Conference on Robotics and Automation (ACRA), Brisbane, Australia, 1-3 December.

ShanieAIBO, *The Sony AIBO Story: From an engineering dream to a robotic pet*, viewed 23 October 2010, <http://www.sonyaibo.net/aibostory.htm>.

Skubic, M, Perzanowski, D, Blisard, S, Schultz, A, Adams, W, Bugajska, M & Brock, D 2004, 'Spatial language for human-robot dialogs', IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 34, no. 2, pp. 154-67.

Sofge, E 2008, *The Problem With Asimo: Tabloid Humanoid Is Holding Back Bots*, 20 May, Popular Mechanics, viewed 23 October 2010, <http://www.popularmechanics.com/technology/engineering/robots/4264593>.s

Steels, L 2001, 'Language games for autonomous robots', *IEEE Intelligent Systems*, vol. 16, no. 5, pp. 16-22.

Steels, L 2008, 'The Symbol Grounding Problem Has Been Solved. So What's Next?', in M de Vega (ed.), *Symbols and Embodiment: Debates on Meaning and Cognition*, Oxford University Press.

Thrun, S, Montemerlo, M, Dahlkamp, H, Stavens, D, Aron, A, Diebel, J, Fong, P, Gale, J, Halpenny, M, Hoffmann, G, Lau, K, Oakley, C, Palatucci, M, Pratt, V, Stang, P, Strohband, S, Dupont, C, Jendrossek, L-E, Koelen, C, Markey, C, Rummel, C, Niekerk, J, Jensen, E, Alessandrini, P, Bradski, G, Davies, B, Ettinger, S, Kaehler, A, Nefian, A & Mahoney, P 2006, ‘Stanley: The robot that won the DARPA Grand Challenge’, *Journal of Robotic Systems*, vol. 23, no. 9, pp. 661-92.

Turing, AM 1950, ‘Computing machinery and intelligence’, *Mind*, vol. 59, pp. 433-60.

Weng, Y-H, Chen, C-H & Su, C-T 2009, ‘Toward the Human-Robot Co-Existence Society: On Safety Intelligence for Next Generation Robots’, *International Journal of Social Robotics*, vol.1, no. 4, pp. 267-82.

Willow Garage, *Willow Garage: PR2 Robotics Platform*, viewed 23 October 2010, <http://www.willowgarage.com/pages/pr2/overview>.

Acknowledgements The author thanks Janet Wiles & Ruth Schulz from the University of Queensland for the research placement opportunity with the Lingodroids Group, as part of the course SCIE1010 Introduction to Research & Practices.
Lingodroids: Language & Robots, and;
Thinking Systems: Navigating Through Real & Conceptual Places;
are projects of the:
UQ CIS (Complex & Intelligent Systems Research Group)
http://www.itee.uq.edu.au/~cis/

Author Information Readers are welcome to comment on this article.
Correspondence should be addressed to D.B. (david.butterworth@uqconnect.edu.au)