

INTELLIGENT ROBOTICS: PAST, PRESENT AND FUTURE

RAY JARVIS

*Director, Intelligent Robotics Research Centre
Monash University, Australia
Ray.Jarvis@eng.monash.edu.au*

This is a position/tutorial paper which begins by tracing the origins and development of Intelligent Robotics and indicates where the field is now and where it may progress to in the near future, covering both fundamental conceptual structures and potential fertile application domains.

1. The Past

Practical robotics had its origins inside factories manufacturing products on assembly lines, where speed, precision, and reliability were paramount. Thus both human tedium and imprecision were done away with whilst improving the uniformity of a quality product which could come off the assembly line relentlessly around the clock. Precision machine tools are considered the inspiration of such a development. However, before robots could be programmed using a computer, it was not possible to easily change any detail of their repetitive function. With the flexibility of programming and the training supplied by a skilled operator, robots could carry out very complex, human-like, repetitive tasks; however the task could be changed at short notice by using a different program. Coping with environmental variability could only be accomplished with sensor feedback, this leading to what was known as the third robot generation. It soon became apparent that higher degrees of intelligence were needed to accommodate variability and uncertainty. As robots migrated out of the fixed automation, fully structured factory assembly lines into the unstructured and unpredictable worlds of space, underwater, in the air and on the ground, where many of the future applications could be imagined, it became clear that a complementary range of sensors and considerable artificial intelligence would be needed to achieve autonomy. This striving for autonomy in complex, unstructured and unpredictable environments, sometimes cohabited by humans became and still is the holly grail of robotics and has given rise to the field of 'intelligent robotics' where perception, reasoning and actuation are highly coupled to achieve useful tasks with little human guidance [See Figure 1]. Pursuit of this goal, whilst promising practical outcomes, is also laced with no small measure of romance, in the idea of fabricating a human like 'creature' of intelligence and grace which can be a human helper and companion.

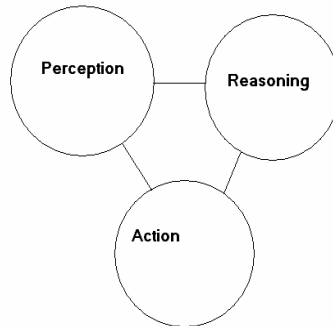


Figure 1. Intelligent Robotics.

3. The Present

Now we have access to a very rich set of sophisticated sensors, powerful computing platforms and all variety of agile mechatronic devices, it would seem that nothing should hold us back in achieving the dream of a fully autonomous robot agent which could carry out a variety of complex tasks in unstructured environments and at the same time be able to interact cooperatively with humans. However this is still far from being achieved. Considerable advances have nevertheless been made.

It is useful to consider the basic ingredients of a fully fledged robotic system at a rudimentary (functionality) level. First, there is a need to know how a robotic manipulator's links are configured in space or where a mobile robot is [Jarvis,1993] and how it is posed. Since manipulator links are interconnected and the geometry is known, a combination of shaft encoder readings (or equivalent) and the derivation of the inverse kinematics can determine the position and pose of the end effector (hand) fairly accurately. Flexure, backlash, and pay-load variations do, however, add error but this is usually manageable, perhaps with the addition of special devices such as lasers and cameras. Attempts to carry out similar evaluations on a free roaming mobile robot using only odometry have not been satisfactory since wheel slippage, load distribution, support surface undulations and wheel shape imperfections introduce errors which are accumulative and unrecoverable from. Following fixed lines on or signals under the floor, whilst workable, restricts the mobility of the robot. The use of beacons at known locations and scanners [See Figure 2] can be very effective but site preparation is still required. There has been a definite trend in recent times to use only natural landmarks to achieve location/pose (localisation) without site preparation, sometimes without even knowledge of the site beforehand. If the site is known, sensors on-board the robot still have to perform the 'data association' task to relate current sensor measurements to previously stored map information. Usually continuity and speed constraints, together with odometric, visual, radar, range (ultrasound or laser based) or inertial information, can simplify this task. However, if the robot is picked up and deposited in a new place without the system's knowledge, determining the robots localisation (the so-called 'kidnapped robot' problem [Spero, 2006] is more difficult and global localisation strategies are required. There are essentially two modalities of natural landmark exploitation for robot localisation, one depending on specific, selected landmarks, and the

other on the entire 'shape' of the sensed environment. Specific landmarks should ideally be unique, detectable, temporally stable, distributed, visually stable (do not look very different from different vantage points) and vertically compact (i.e. have a clear X, Y position on the terrain rather than spread out like a lake). When comparing live sensor data acquired on-board with a known environment map, the landmark correspondence matching problem must be solved with minimal ambiguity and some degree of redundancy for accuracy and reliability. Mismatches can have serious localisation consequences but continuity and velocity constraints can be used to detect inconsistencies or reduce ambiguity to an acceptable level. The second modality consists of using the entire set of live sensor data; preferably dense (as for vision and high resolution laser range scans) to estimate where the robot must be to make that observation by matching against the environmental map database. Pure appearance based or range/image based sensor data [Jarvis *et al.*,2007] is usually employed.



Figure 2. Bar Code Scan Localiser.

A large variety of matching techniques have appeared in the literature. The application of particle filters [Fox *et al.*,2001] which is a form of genetic search algorithm where randomly selected hypothesised locations are tested against a matching criterion and the probabilistic search (by resampling) narrows in favour of the better matches until a sufficiently compact distribution of feasible results is obtained (the centroid of which may be used as the safest sought-after position/pose). Odometric and inertial sensor data can be used to predict relative localisation to allow the drift of 'particles' to more likely hypothesis positions and reduce the computational search costs.

Attempts to localise without a prior environmental map would seem to present a dilemma. From a known initial position a robot can collect environmental sensor data and save a map relative to its position/pose. If, when the robot moves it collects new data, then matching this against the already saved environmental data (with which a reasonable overlap is maintained), the new position/pose can be estimated but some error will be introduced. The new environmental data beyond the overlap can now only be saved with some error. As the robot can only determine its moving position by matching against faulty data, an accumulation of both localisation and mapping error will result, eventually leading to very distorted notions of where the robot is and how the environment is structured.

This has become known as the 'simultaneous mapping and localisation' (SLAM) problem [Dissanayaka *et al.*,2001] and has been a centrepiece of effort amongst the robotics research community for over 15 years. If the original position/pose can be gained with certainty, error corrections can be distributed around all the route points and the environmental map corrected to a large extent. This closure requirement has also been a focus of concern. By using Kalman predictive filters [Kelly,1994] and particle filters, optimal strategies for solving the SLAM problem have been proposed. Yet the data correspondence and closure problems are still subject to further refinements and extensions using multiple sensors (sensor fusion) including cameras laser range sensors, odometry, radar ultrasound and inertial sensing. In outdoor situations a Global Positioning System (GPS) can provide global localisation where the satellite signals are received adequately but there are many situations (built up areas, dense forests, deep valleys, underground, underwater etc.) where GPS can not be deployed and the sophisticated SLAM methods are required where map data is not already available. Probabilistic SLAM methodology [Thrun *et al.*,1998] is very popular currently.

If an environmental map is not provided initially, finding ways of constructing such a map using the robot's sensors is of considerable interest in its own right in addition to providing the bases of natural landmark localisation as described above. Over the last 20 years or so a large variety of sensor technology has become available to achieve this goal. Since the mobile robot's workspace is essential 3D, what is required is essentially a volumetric map of occupied space, possibly colour rendered [See Figure 3]. Two or more video cameras with baseline separations can be used to achieve stereovision, extracting depth information by evaluating the disparities amongst the images in relation to target points in 3D space. Again the data correspondence problem arises and only unique and correct matches will produce valid data. Usually area correlation or edge matching is used. Using more than a minimum of two cameras can help reduce the ambiguities but some visual texture is needed for the matching task, so visually bland regions cannot be ranged to in this way. Range accuracy diminishes with actual range and parts of the scene not viewable by at least two cameras cannot be ranged to. In practical terms such systems are not useful for extracting range beyond approximately 4 metres with camera separations of 15cms. The nature of correspondence matching is still central when panoramic mirror systems [Chal and Srinivasan,1997] or wide base line methodologies such as Scale-Invariant Feature Transform (SIFT) [Lowe,2004] are used. The properties of intensity gradients surrounding salient points are used as signatures to be matched, with fine transformation parameters extracted as results. A single camera moving in the environment can also return a type of stereo data known as 'optical flow'. Panoramic systems have the advantage of seeing all around the robot including where it is going and where it has come from and this can be an advantage for localisation as it avoids the 'window' problem of having perhaps partially overlapping framed views of the world to deal with.

Time-of-flight range sensors such as ultrasonic or laser devices collect data along the same axis as the outward beam (approximately) and thus do not have a base line problem, but do nevertheless have specular reflection difficulties, possibly leading to false or no readings. Ultrasound also has a beam width problem unless special concentration techniques (eg. multiple transducers with main beam alignment) and safer

severe signal attenuation problems in air. They are useful only up to about 15 metres. Laser range finders also suffer from specularly but can range over many thousands of metres. Common laser range finders found in robotics laboratories [See Figure 3.] include Erwin Sick, Hokuyu and Riegl. Riegl laser range scanners can be fitted with a high-resolution digital camera and produce colour rendered 3D maps with ranges up to 800m at sub-centimetre accuracy. Of course only the surfaces visible from the sensor can be scanned and multiple scans from different vantage points are needed to get a full 3D map (a single vantage point map is often called 2½ D).



Figure 3. Riegl Laser Scan of Lab.

Once the localisation and mapping aspect are resolved there still remains the dual requirements of path planning and obstacle avoidance for a mobile robot to navigate from a starting position to a nominated goal point via an efficient collision-free path. There are a large number of path planners described in the literature. A representable set is the A* methodology [Lozano-Perez and Wesley,1979], Rapidly Exploring Random Trees (RRT) [LaValle and Kuffner,2001] and Distance Transforms(DT) [Jarvis,1994]. Each has its own advantages and disadvantages. The distance transform has the advantage of accommodating initially unknown environments, time-varying environments, variations in terrain navigability, explorations of unknown spaces, covering all accessible territory and planning overt paths hidden (or minimally visible) from certain vantage points. It can also be easily extended into any number of dimensions. As it operates in tessellated spaces, it tends to be memory hungry when high resolutions of large spaces are to be covered. The A* and RRT methodologies have their own drawbacks but are applicable to real Euclidean spaces and are more geometric than area based. Some examples are shown in Figure 4.

The matter of locomotion is solved by deciding whether wheels, legs, tracks or whatever means of confronting the terrain suits the application. Legged motion has always been of interest because of biological correlates and dextrous terrain coverage and climbing capabilities.

The distribution of computational and communications resources is a matter of preference in terms of speed, reliability convenience and costs. The actual function of the robot is often neglected in the process of designing the navigation system. Whether the machine is to put out fires, rescue people discover buried mines or carry tree trunks should of course be considered early on not just when everything else has been decided.

Over the last 12 years or so, since Honda announced its P2 bipedal walking humanoid robot, there has been a great deal of research effort but into this particular type of machine. The humanoid form [See Figure 5.] has been much in the forefront of robotics research partly because of the technical challenge it represents, partly because that form can accommodate to human activities in a human centred environment and partly because of the romanticised notion that the human form is worth emulating in its own right and the fascination associated with the possibility of building a truly autonomous human-like creature. The considerable effort in constructing a bipedal gait machine capable of very rough terrain navigation, climbing and flexing may, in the longer term, be considered unjustified, but the challenge itself is real, like putting a man on the moon in the 1960's. Similarly, considerable spin-offs can be predicted.

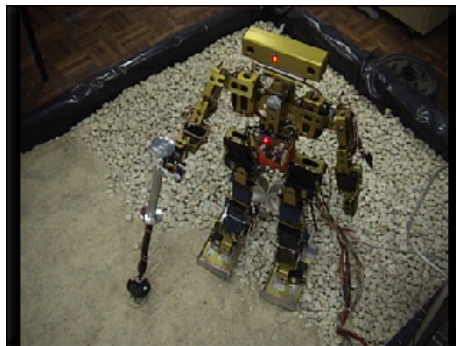


Figure 5. Bipedal Gait Humanoid

Whether human-like or not, a robot may need to communicate and cooperate with humans. This issue has loomed as a difficult challenge, which must be met if human/robot interaction is to become natural and useful. This is particularly true in assistive technology applications where a robot may be required to assist aged, frail, or otherwise disabled citizens in domestic or institutional domains [See Figure 6].



Figure 6. Wakamaru Service Robots.

In recent times the function of speech, gesture and gaze in disambiguating user intent in instructing an assistant robot has become of particular research interest. The robot needs both spatial intelligence to navigate its environment and find things but also needs transactional intelligence to understand and refine the user's requests and negotiate an action within the realms of possibility. One example might be where the user has requested a red mug be brought from the kitchen table but only a blue one can be found there. Perhaps the blue one will do or the red one may be elsewhere and needs to be found. Face recognition may also be important to identify the user and accommodate individual traits, identifying a visitor or at least know to whom a requested item should be brought if that person has moved since the request was made.

Another area of robotics gaining interest of the late is the deployment of swarms [See Figure 7.] of cooperating robotic agents with various collective intelligence strategies, to carry out certain tasks more effectively than possible with one, perhaps more intelligent but cumbersome machine. Often these studies are inspired by the known behaviours of social insects and the graceful degradation of functionally possible with an appropriate swarm design. A fascinating concept named 'emergent behaviours' where complex collective behaviours is manifested by swarms where each individual is guided by simple rules, has been imbued with a certain degree of mysticism. Yet it can be demonstrated that certain emergent behaviours can be designed [Lai *et al.*2007]



Figure 7. Robot Swarm.

Although it has been a long time coming, the age where robots of various kinds and capabilities will work along-side humans [Harte,2007] or in hazardous environments where humans dare not venture is soon to be realised. The long delay in the realisation of this possibly has at least three contributors. The first has been the lack, until recently, of relatively inexpensive sensor and computational technology. The second is the simple fact that many activities easily carried out by humans are much harder to realise in a machine than was originally thought. Huge amounts of effort in devising artificial intelligence and signal processing methodologies has only recently yielded modest commercially applicable results. Safety, reliability, and cost have been largely neglected until recently. The third factor is probably that the need was not sufficient to push robotics technology into the market place and the cost-justifiable application scope of robotic deployment was not well enough articulated. Of course there were always people looking for smart robotic gadgets but not that many say compared to computer communications and entertainment electronics enthusiasts.

4. The Future

Leaving Science Fiction aside, the expectations concerning intelligent robotic technology development over the next decade or so are quite modest. The practical application domains where robotic technology is most likely to be used are:

1. Transport (public and private)
2. Exploration (oceans, space, deserts etc.)
3. Mining (dangerous environments)
4. Civil Defence (search and rescue, fire fighting etc.)
5. Security/Surveillance (patrol, observation and intervention)
6. Domestic Services (cleaning etc.)
7. Entertainment (robotic toys etc.)
8. Assistive Technologies (support for the fragile)
9. War Machines
10. Scientific Instrumentation (e.g. synchrotron sample preparation, chemical screening etc.)

Since the last of these is more like an assembly line process, it has been one of the first to have already been deployed commercially. Yet complex aspects, such as the capture and mounting crystals for x-ray crystallographic analysis, have not yet been automated.

All the other domains involve high degrees of unstructuredness, where commensurate levels of artificial intelligence based on sensor data fusion and understanding are required.

An interesting way to approach autonomous operation whilst realizing useful applications along the way is to devise the means by which humans can interact and intervene [See Figure 8.] with robots which are richly sensor equipped, providing the missing capabilities such as subtle judgments, risk analysis, fine dexterity and reaction to unpredicted events but in such a way that a continuum between full autonomy and full teleoperation can be smoothly engaged. As the technology matures, the human support can be gracefully withdrawn with less and less intervention over more and more of the tasks. For example there are many situations in, say, fire fighting where a robotic vehicle carrying water may move along a fire front spraying water at hot spots detected using a thermal camera fairly autonomously, but a human may need to direct the vehicle to move to another more critical, fire front or assist in a delicate rescue mission under direct teleoperational control. As another example, a transport vehicle may safely negotiate a highway navigation task, changing lanes and speed as required, replanning routes and so on, yet a human operator may need to take over at an unexpected construction site or scene of an accident.



Figure 8. Semi-Autonomous Wheelchair.

Some new robotics research interests are developing in the bio-robotics and cognitive science fields, where the melding of biology with robotics has exposed new

insights into the functions of mammalian brains, pathways to muscle functions and perception and has led the way towards evolving new prosthesis devices for both mechanical and perceptive needs. Whilst biologists and cognitive scientists are using robotic notions to help them develop models of physiological function and test them in robotic experimental ways, so also have roboticists been inspired by deeper understandings of biological processes in their design of robots [Beer et al.,1997].

Thus the future for robots being used at various and variable levels of autonomy in many everyday activities is likely to be bright if one does not expect too much. The more exciting challenges will be where humans and robots work together cooperatively with a need for reliable communication and due regard for safety and reliability. Probably only mass production will permit the domestic robots to be affordable by most in much the same way as automobiles have proliferated.

It is likely that affordable computational power will continue to grow, perhaps doubling in processing capability every two years at the same cost as it has for some time already. Sensor technology will not only continue to improve in capability but its price will rapidly decline. Initial navigation systems which cost tens of thousands of dollars a decade ago are now available for a few thousand dollars or less. The development of robust methodologies for navigation, recognition, and human interaction will continue, with modest improvements generated incrementally rather than in leaps and bounds. It is clearly here where the most research effort is needed. The melding of computational and robotic sciences will be essential for these improvements.

One somewhat frightening possibility may impinge upon the development of personal robots to do chores for humans. The fear that genetic engineered sub-humans may take over such roles is certainly worth contemplating but preferably not embracing. Imagine a world in which being much like ourselves are happy to do our chores for very little reward or freedom. Where then would robotics still be relevant? That such developments are currently regarded as unethical and immoral may not be sufficient deterrent for such developments, perhaps initially by the unscrupulous, but may be later adopted more generally. Perhaps mere ignorance has raised this concern. Surely heavy duty operations in mines and industry, mineral exploration, hazardous environments and a host of other applications of this kind would still be robot related, but the need for sophisticated machine intelligence could be severely diminished.

The main challenges for the future of intelligent robotics are:

1. Improvements of the quality, robustness, smaller size and reduced cost of camera, laser range, ultrasonics, radar, and inertial sensors.
2. Improvements in computational power at low cost. This aspect will not need any special attention because of existing market forces.
3. Improvement in mechanisms for robot platforms in terms of weight, strength, and capability and the use of new materials, including ceramics, carbon fibre, titanium etc.

4. Improvements in navigation algorithms including natural landmark based approaches, recovery mechanisms, accommodation of varying cost structures related to navigability, collision risk, visibility etc.
5. Improvements of Human/Machine cooperation, including communication, task refinement, intervention etc.
6. Improvement in risk assessment and endurance in terms of operational times and graceful degradation.
7. Clarification of legal aspects of humans and robots working together.
8. Better understanding of emotional aspects of robots working with humans.
9. Evolution of the robot/biology cross-inspirational trend.
10. Development of robotic ethics.

Conclusions

The overall picture for the development of intelligent robotics is very positive but more evolutionary than revolutionary, with a steady penetration into the industrial and domestic worlds at affordable prices on the near horizon. It would be expected that, as the price per unit reduces through mass production as it has for automobiles and personal computers, there will become a time when simple robotic devices for use in the home will become standard peripherals (like printers) and many consumer items from washing machines and lawn mowers, to automobiles and entertainments systems will have intelligent robotics capabilities built in. This process will start as a trickle but soon move to deluge proportions, probably without great acclaim, as has been the case when technologies once regarded as mind blowing become commonplace and merely a matter of affordability and compliance with legal systems. The substitution of sensor based intelligent for speed and precision to accommodate uncertainty and unstructuredness will remain the main driving force, being the essential paradigm shift.

References

- Beer, Randall D., Roger D. Quinn, Hillel J. Chiel, and Roy E. Ritzmann.(1997) [Biologically Inspired Approaches to Robotics](#). Communications of the ACM, March : 30-38.
- Chahl , J.S.and Srinivasan, M.V.,(1997) Reflective surfaces for panoramic imaging. *Applied Optics* 36, 8275-8285.
- Dissanayaka, M., Newman, P., Clark, S., Durrant-Whyte, H., and Csorba, M.(2001) A solution to the simultaneous localisation and map building problem. *IEEE Trans. on Robotics and Automation* 17, 3 June , 229-241.
- Fox, D., Thrun, S., Burgard, W, and Dellaert, F.(2001) [Particle filters for mobile robot localization](#). In Doucet, A., de Freitas, N. and Gordon, N. editors, *Sequential Monte Carlo Methods in Practice*, pages 499-516. Springer Verlag.
- Harte, E. and Jarvis, R.A.,(2007) Multimodal Human-Robot Interaction in an Assistive Technology Context using Environmental Knowledge., accepted for presentation at

- the Australasian Conference on Robotics and Automation, 10th to 12th. Dec, Brisbane, Australia.
- Jarvis, R.A.(1993) A Selective Survey of Localisation Methodology for Autonomous Mobile Robot Navigation, accepted for presentation at the Robots for Competitive Industry Conference, July 14-16, Brisbane Australia, pp. 310-317.
- Jarvis, R.A.(1994) On Distance Transform Based Collision-Free Path Planning for Robot Navigation in Known, Unknown and Time-Varying Environments, invited chapter for a book entitled 'Advanced Mobile Robots' edited by Professor Yuan F. Zang World Scientific Publishing Co. Pty. Lt, pp. 3-31.
- Jarvis, R.A., Ho, N. and Byrne, J.B.(2007) Autonomous Robot navigation in Cyber and Real Worlds, CyberWorlds 2007, Hanover, Germany, Oct. 24th to 27th, pp. 66-73.
- Kelly, A.(1994) A 3D Space Formulation of a Navigation Kalman Filter for Autonomous Vehicles, tech. report CMU-RI-TR-94-19, Robotics Institute, Carnegie Mellon University, May.
- Lai, L., Manning, J, Su, J., Kazadi, S.,(2007) A Principled Approach to Swarm-Based Wall-Building, Progress in Artificial Life, Third Australian Conference, ACAL 2007, Gold Coast, Queensland, Dec., pp305-319.
- LaValle, S.M. and Kuffner.(2001), J.J, Rapidly-exploring random trees: Progress and prospects. In B. R. Donald, K. M. Lynch, and D. Rus, editors, *Algorithmic and Computational Robotics: New Directions*, pages 293--308. A K Peters, Wellesley, MA,
- Lowe, D. G.,(2004) "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, 60, 2, pp. 91-110.
- Lozano-Pérez, T., and Wesley, M.A.,(1979) An Algorithm for Planning Collision-Free Paths Among Polyhedral Obstacles. *Commun. ACM* 22(10): 560-570
- Spero, D., (2006) Simultaneous Localisation and Map Building: The Kidnapped Way.
- Thrun, S., Burgard, W., and Fox, D. A (1998) probabilistic approach to concurrent mapping and localization for mobile robots. *Machine Learning* 31 , 29-53.