appeared 4th Jul 2007

Keen on the scent

Getting good results from uncertain data is important for getting on in the wild as well as in the marketplace, says S.Ananthanaryanan.

Nature has carried a report of a mathematical approach to a version of the problem, which has led to a solution that is akin to the method used by the common moth in seeking a mate in windy conditions.

Following scents

In the child's game of finding a hidden object, players are guided by calls of 'hot', 'warm', 'cold', 'cooler', as the players get closer or further from the target. Bacteria seeking nutrients or a dog following a burglar do something similar – they take the direction along which the scent keeps getting stronger.

This method works so long as the scent is strong enough to show a gradient and also has not been dispersed or broken by cross winds. But more complex methods are necessary when far away from the source or when the signal is garbled or intermittent.

Signal and noise

When the signal gets weak or broken, the problem of misleading, 'accidental' smells or traces, which mask the real trail, becomes important. It is like holding a conversation in a noisy cocktail party – it is difficult to hear what your partner is saying because of the many voices speaking at once. The partner's words, the 'signal', are drowned in the 'noise', or sounds that are not 'signal'.

The communications engineer Claude Shannon invented the science of information theory and developed an exact formula for the maximum information that can be carried over a channel (like a telephone wire) which is subject to 'noise' (like random electrical pulses due to thermal excitation of the atoms of the wire). Shannon's monumental work has done much to help bring about the revolution in electronics and communications in the last half century.

The effect when the signal is weak or scattered is that the proportion of real signal to 'noise' reduces and so does the 'information'. The need then is to 'listen' for longer, to collect more information. The problem is that when the signal reduces very fast, the waiting time till something intelligible and reliable is at hand may become prohibitive. The need then is to 'optimise', for a compromise between reliability and speed.

Search algorithm

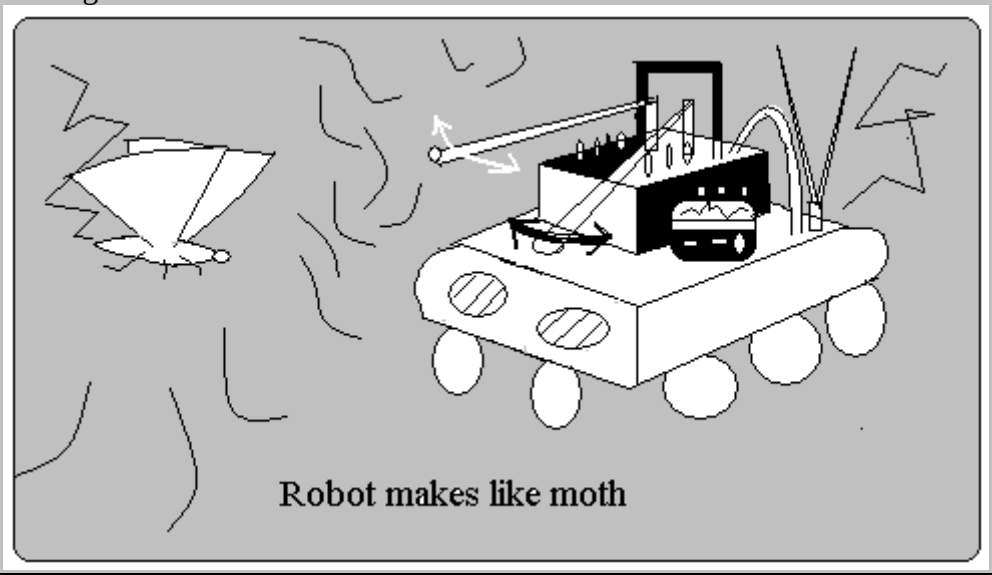
When the signal beyond the immediate vicinity does not show a continuous gradient, or increase along some direction, a strategy has to be of "polling" the surroundings for more information. Vergassola, Villermaux and Shraiman in ***Nature*** last week reported a system of dong this, where the direction to follow is indicated by the direction in which the rate of information gain is maximum.

In information theory, information is defined as a reduction of uncertainty. Thus, to resolve the correct choice out of a field of two possibilities is to have one unit of information, or '1 bit'. And it takes 8 bits of information to resolved 256 (or 2 to the power 8) possibilities. These principles, along with statistical methods are able to take a weak and uncertain sample and generate a field of possible sources of signals. The weaker the sample the more blurred the picture of the source.

In the method developed by the authors of the paper, the recipient does not sit still to wait for information to collect, but moves rapidly around, picking up data of the distribution of weak or strong or no signal. As the data builds, the image of the source gets sharper and this, in turn, directs the staccato data gathering – till the signal is strong enough to follow directly. The zigzag phase, called 'infotaxis' relies on maximising 'information gain', just as following the direct path, or 'chemotaxis', is to maximize the strength of the signal itself.

Moths and robots

The way a king moth can follow the pheromone trail to a female kilometres away has been studied and documented. At great distances the trail is both almost the same, low strength everywhere as well as liberally interspersed with 'no trail'. The moth characteristically follows a zigzag pursuit or 'casting' to search for a trail when lost, to drift in the right direction.



Tracing the source of leaks of poisonous or inflammable gasses, or contaminants, is of importance in industry and in environmental or demographic work. The algorithm developed could help robots and researchers alike, in rapidly tracing the source of signals, based on weak samples.