

## ***Homer (by Vere & Bickmore) as a Conscious Agent***

Ref: S. Vere and T. Bickmore, "A basic agent", Computational Intelligence 6(1), February 1990, pages 41-60, <https://doi.org/10.1111/j.1467-8640.1990.tb00128.x>

This is a landmark paper in that it combines multiple human-like abilities (in a very limited way of course) in a single agent, Homer -- dialogue capabilities, time-dependent planning, navigation and other actions in accord with user-requested goals, and simulated perception (detecting objects in a fan-shaped region towards its front).

The only comparable precursor was Winograd's SHRDLU system for a simulated blocks world (e.g., see <http://hci.stanford.edu/~winograd/shrdlu/index.html>), and there have been relatively few comparably ambitious & versatile systems since. (There is a brief survey of such work in the paper by Liu & Schubert on the Gridworld framework, which we will also look at closely.)

### ***Is Homer a conscious agent?***

That's the question we want to focus on here, by considering Homer's abilities, how they are implemented, and how those abilities and the way they are implemented relate to consciousness ("C").

Of course, we know the standard distinction between "access C" and "phenomenal C", and our focus will be on access C -- partly because ascribing phenomenal C to Homer seems far-fetched, and partly because that is a topic we will get to separately later.

So let's enumerate Homer's abilities, consider how they are implemented, and try to understand their relevance (or irrelevance) to access C -- in particular self-awareness.

### ***Homer's world, Homer's abilities, and their implementation***

1. First, we should consider "Seaworld", an interesting but still very limited 2-D environment wherein Homer -- depicted as a small submarine-like agent -- lives and acts. There are about two dozen objects, mostly of different types. While various objects, such as the tokens for "Steve" and "Tim", logs, barges, birds, etc. can be "dragged" to different places by the user, only the "Smirnoff" submarine is spontaneously active, patrolling back and forth. What does Homer *know* about its world? Well, it knows locations of things it has encountered or been told about, and it also has a lot of "ISA" knowledge, such as what entities are persons, animals, objects, or locations. It also has a very good understanding of time -- an important aspect of an agent that "lives" in time and space. It also can be given facts such as what things are alive, and their degree of emotional activation.

Is there more? We should note that Homer is also linked to the real world, *via its linguistic user interface*. This *potentially* expands the scope of its knowledge and interaction. However, while Homer's knowledge can be expanded by telling it "facts" such as who owns what object, and where objects are located, it was not intended to reason about the real world beyond "Seaworld".

*Does it make a difference in assessing the C of an agent, specifically its self-awareness, how complex its world and world knowledge are? We can probably agree that self-awareness (and C more generally) is*

a matter of degree: There just isn't much self-knowledge an agent can have in a world where there are few entities and the scope for action is very limited!

2. Second, Homer is able to perform a variety of actions in its world, including (a) travelling to various destinations, while avoiding obstacles, (b) "taking pictures" of encountered objects if asked to do so, (c) "seeing" and identifying objects that come into its fan-shaped field of view, and (d) picking up and delivering certain objects such as packages. These actions are implemented in triple models:
  - the first type of model, which is similar to the STRIPS approach, is intended to enable language understanding;
  - the second is an "expansion" of the linguistic models, with such details as fuel consumption and timing details, not included in the linguistic models;
  - the third is the code used for the actual simulation of actions, which also uses a more or less declarative language, but with various complications, such as separation of actions into phases, formulas for computing quantities, and directives about goal expansion.

The first two types of models use ANTECEDENT and CONSEQUENT sections that are much like *preconditions and effects* in STRIPS-like models. However, there are complications due to the fact that

- a specific action such as **go** can be specified in terms of a more general action such as **move**, which can in turn be specified in terms of a **state change**;
- these various actions have constraints, attributes, and word-specific information wrapped around them, to link up with English; e.g., that "go" is intransitive, what its irregular forms (gone, went) are, and what prepositional forms can be used with it. Constraints are ones like (where variables are preceded by dots)  
(ISA .agent ANIMAL), (NOT (ISA .old.place .new.place)).

*If an agent has very limited, or no, capacity for action, can it still have a high degree of self-awareness?* In class discussion, the view emerged that as long as thought and at least inner language exist, there could still be full self-awareness – as in a fully paralyzed but conscious person.

3. Third, Homer has a sophisticated temporal planning capability, deriving from Vere's work on temporal planning for a spacecraft sending pictures of planets back to Earth. This planning capability works much like what we've discussed about STRIPS-based planning, except for the complications involved in dealing with time and other quantities.

*Can an agent without a planning capacity still have a high degree of self-awareness?* To the extent that planning is the hallmark of human intelligence, human-like self-awareness really implies coupling of access to self-knowledge to a sophisticated planning capability. However, this need not necessarily involve physical action, but might be limited to the unconscious and to some degree conscious ability to plan *speech acts* and extended *discourse* – which after all have a coherent and often purposeful structure, and can powerfully influence individuals and the world (as we can see in the movie *Her*, or in current politics).

4. Fourth, and equally impressively, Homer has quite a sophisticated natural language dialogue ability, aimed at enabling it to plan for, and carry out requested actions, and answer questions not only about the present situation, but also about past events (*episodic memory*) and expected future situations, given its current plans.

*This ability seems particularly crucial for a high level of self-awareness* – there seems to be a tight link between what we are aware of, and what we can report (or at least formulate inwardly) in language.

As noted in the discussion, a person with Alzheimers whose gradually loses their episodic memory and the ability to converse seems to have ever less self-awareness.

- 5 . Fifth, there is a limited inference capability, e.g., inferring that an object has “disappeared” (if it was seen in a place before and was no longer there later), or has been “passed” (if it was seen up ahead before, and Homer went forward further than the original distance of the object). But perhaps more importantly, we should reckon Homer’s planning ability as a form of inference; recall that planning can be viewed as a constructive sort of inference; more specifically it may be viewed as *abductive* inference or even *deductive* inference (proving that there exists a sequence of actions that gets the agent to the goal state)

*Is inference essential for self-awareness?* Well, humans have vast amounts of inference-enabling knowledge about the world – about thousands of different types entities, their properties, behaviors, and uses, and about humans in particular – especially a “theory of mind” that enables inference of what people intend, think, and feel under various circumstances. This knowledge enables commonsense inferences about future events and behaviors (the sun will rise, the red light will turn green, the flight I’m on won’t crash, if I go to a restaurant I’ll be able to order and eat a meal, my friends will help me if I need help, if I help my friends they’ll be appreciative, children want to play and explore and be taken care of, if I crash my car it’ll be wrecked and I may be hurt and land in the hospital or worse, etc.) To the extent that self-awareness requires knowing not only one’s own properties but also one’s interactions with the world at large, inference based on knowledge is important for attaining a high degree of self-awareness. Homer has sophisticated planning abilities, but they are limited to a small, sparse world (e.g., there is no “theory of mind”), and so it has only quite limited self-awareness.

- 6 . Finally, there is “visual perception”, but that is limited to detecting objects that fall within its forward-directed, fan-shaped field of vision. In addition, Homer “hears” what the user says to it.

*Is perception essential to self-awareness?* Well, Helen Keller had no vision or hearing, but her sense of touch (along with smell, taste and other bodily senses) opened a path to a full human, indeed distinguished, life. It seems that as long as there are “channels to the world” of some sort, full self-awareness is not necessarily impeded. However, phenomenal consciousness is another matter. Homer probably has no significant degree of that ...