computers are bad

https://computer.rip - me@computer.rip - fax: +1 (505) 926-5492

2022-11-23 enlightenment and lighting controls

One of my chief interests is lighting. This manifests primarily as no end of tinkering with inexpensive consumer IoT devices, because I am cheap and running new cabling is time consuming. I did nearly end up using DMX for my under-cabinet lighting but ultimately saw sense and stuck to a protocol that is even more unfamiliar to the average consumer, Z-Wave.

I worked in theater (at a university conference center) only briefly but the fact that it was a very small operation gave me a great deal of exposure to the cutting edge of theatrical control last time a major capital expenditure had been authorized, in the 90s. This was an ETC Sensor dimmer system with an ETC Express 48/96 console for which we had to maintain a small stash of 3.5 diskettes. The ETC Express is still, in my mind, pretty much the pinnacle of user interface design: it had delightfully tactile mechanical buttons that you pressed according to a scheme that was somehow simultaneously intuitive and utterly inscrutable. Mastery of the Thru, And, Except, Rel buttons made you feel like a wizard even though you were essentially typing very elementary sentences. It ran some type of non-Microsoft commercial DOS, maybe DR-DOS if I remember correctly, and drove the attached 1080p LCD display at 1024x768.

The integration with the Lutron architectural lighting control system had never actually worked properly, necessitating a somewhat complex pattern of button-mashing to turn on the lobby lights that sometimes turned into sending a runner upstairs to mash other buttons on a different panel. There was an accessory, the Remote Focus Unit, that was a much smaller version of the console that was even more inscrutable to use, and that you would carry around with you trailing a thick cable as you navigated the catwalks. This was one of two XLR cables that clattered along the steel grating behind you, the other being for the wired intercom system.

My brief career in theater was very influential on me: it was a sort of Battlestar Galactica-esque world in which every piece of technology was from the late 90s or early 00s, and nothing was wireless. You unplugged your intercom pack, climbed the spiral staircase (which claimed many a shin) to an alarmingly high point in the flyloft, and plugged your intercom pack into the wall socket up there. Then you fiddled around for a moment and had to walk back to the wall socket, because the toggle switch that changed the socket between buses was always set wrong, and you never thought to check it in the first place. Truly a wonderful era of technology.

The spiral staircase exists in a strange liminal space in the building: the large open area, behind the flyloft, which primarily contained the air handlers for the passive solar heating system installed as a pilot project in the 80s. It had apparently never worked well. The water tubing was prone to leaking, and the storage closets under the solar array had to be treated as if they were outdoors. Many of the counterweights and older fixtures were rusty for this reason. It would rain indoors, in the back of the Macey Center: not because of some microclimate phenomena, but by the simple logic of a university that occasionally received generous grants for new technology, but never had the money for maintenance. Of course today, the passive solar array has been removed and replaced by a pointless bank of multicolored architectural panels curiously aimed at the sun. Progress marches on.

Well thats enough nostalgia. Heres the point: I think lighting control is interesting, chiefly because it involves a whole lot of low-speed digital protocols that are all more or less related to RS-422. But also, there is history!

I am going to sort of mix theatrical and architectural lighting control here, but it is useful to know the difference. Theatrical lighting control is generally the older field. Early theaters had used chemical light sources (literal limelight) to light the stage, later theaters used primitive electrical lights like carbon-arc spotlights. Just about as soon as electrical lights were introduced, electrical light *dimming* came about. Theatrical lighting controls have certain properties that are different from other lighting systems. They are usually designed for flexibility, to allow light fixtures (sometimes called luminaires but moreso in architectural lighting) to be moved around and swapped out between shows. They are designed with the expectation that relatively complex scenes will be composed, with a single show containing a large a number of lighting cues that will be changed from show to show. Theatrical lighting is largely confined to the theater, mostly in very traditional forms of footlights, side lights, and numbered catwalks or bridges extending both directions from the proscenium (upstage and into the house).

Architectural lighting systems, on the other hand, are intended to make buildings both more dramatic and practical. There are similarities in that architectural lighting control systems mostly involve channels which are dimmed. But there are significant differences: architectural lighting is mostly permanently installed and unmovable. There is a relatively limited number of channels, and more significantly there is a relatively limited number of scenes: maybe a half dozen in total. Control is sometimes automated (based on a solar calendar, sunset and sunrise) and when manual is intended to be operated by untrained persons, and so usually limited to a row of buttons that call up different scenes. You, the reader, probably encounter architectural lighting control most often in the completely automated, scheduled systems used by large buildings, and second in the wall panel scene control systems often used in conference rooms and lecture halls.

There exists, of course, an uncomfortable in between: the corporate or university auditorium, which has some elements of both. Theaters are also often found inside of buildings with architectural lighting controls, leading to a need to make the two interoperate. Because of both the similarities and the need for interoperability there are some common protocols between theatrical and architectural systems, but for the most part they are still fairly separate from each other.

So how does a light control system actually work?

The primitive element of a light control system was, for a long time, the dimmer. Early theaters used saltwater dimmers and later variac dimmers arranged into banks and operated by levers, which could be mechanically linked to each other to effect scene changes. Architectural systems are much the same, but instead of backstage or in a patch bay, the dimmers are located in a closet. Architectural systems have always been more automated and required remote control, which of course means that they came about later.

Lets start with a very basic but very common scheme for lighting control: the 0-10V dimmer. Widely used in architectural lighting for many decades, this is perhaps the simplest viable system. For each dimmer there is a knob which adjusts an output voltage between 0 and 10v, and this is routed by low voltage wiring to either a central dimmer or (more common in later systems) a distributed system of dimmmable lighting ballasts incorporated into the fixtures. The main appeal of 0-10v analog dimming is its simplicity, but this simplicity betrays the basic complexity of dimming.

Some lights are very easy to dim, mostly incandescent bulbs which are capable of a very wide range of brightnesses corresponding more or less linearly to the power they consume (which can be moderated by the voltage applied or by other means). Arc and discharge lights introduce a complication; they produce no light at all until they reach a striking power at which point they can be dimmed back down to a lower power level. Incandescent light bulbs can actually behave the same way, although it tends to be less obvious. The issue is a bigger one in architectural lighting than in theatrical lighting [1], because architectural lighting of the early era of central control relied heavily on fluorescent fixtures. These have a particularly dramatic difference between striking power and minimum power, and in general are difficult to dim [2].

This has lead to a few different variations on the 0-10v scheme, the most common of which is 1-10v fluorescent control. In this variant, 0v means off while 1v means minimum brightness. This difference is purely semantic in the case of incandescent bulbs, but for fluorescent ballasts indicates whether or not the bulb should be struck. The clear differentiation between off and very dim was important for simpler, non-microcontroller ballasts, but then became less important over time as most fluorescent ballasts switched to computerization which could more intelligently make a threshold decision about whether or not the bulb should be struck near 0v.

The 0-10v scheme is simple and easy to work with, so it was widely installed. It has a major downside, though: the need to run separate control wiring to every zone or set of dimmable lights. In typical architectural installations this is pretty manageable, because in the era of 0-10v analog dimming (as in the era before of direct dimming of the power supply wiring) it was typical to have perhaps six distinct zones. In theatrical lighting, where a modest configuration is more like 16 dimming channels and wiring is more often expected to be portable and reconfigurable, it was a much bigger nuisance. Fortunately, improving electronic technology coming mostly out of the telecom industry offered a promising innovation: multiplexing.

If you are not familiar with the term at its most general, multiplexing describes basically any method of encoding more than one logical channel over a single physical channel. On this blog I have spoken about various multiplexing methods since it has an extensive history in telecommunications, most obviously for the purpose of putting multiple telephone calls over one set of wires. If you, like me, have an academic education in computing you might remember the high level principles from a data communications or networking class. The most common forms of multiplexing are FDM and TDM, or frequency division muxing and time division muxing. While Im omitting perhaps a bit of nuance, it is mostly safe to say that muxing is an abbreviation for multiplexing which is the kind of word that you quickly get tired of typing.

While there are forms of muxing other than FDM and TDM, if you understand FDM and TDM you can interpret most other methods that exist as being some sort of variation on one, the other, or both at the same time. FDM, frequency division, is best explained (at least I think) in the example of analog telephone muxing. Humans can hear roughly from 20-20kHz, and speech occurs mostly at the bottom end of this range, from 80-8kHz (these rough ranges tend to keep to multiples of ten like this because, well, its convenient, and also tends to reflect reality well since humans interpret sound mostly on a logarithmic basis). A well-conditioned telephone pair can carry frequencies up to a couple hundred kHz, which means that when youre carrying a single voice conversation theres a lot of wasted headroom in the high frequencies, higher than audible to humans. You can take advantage of this by mixing a speech channel with a higher frequency carrier, say 40kHz, and mixing the result with an unmodified voice channel. You now have two voice conversations on the same wire: one at 0-20kHz (often called AF or audio frequency since its what we can directly hear) and another at 40-60kHz. Of course the higher frequency conversation needs to be shifted back down at the other end, but you can see the idea: we can take advantage of the wide bandwidth of the physical channel to stuff two different logical channels onto it at the same time. And this is, of course, fundamentally how radio and all other RF communications media work.

TDM, time division, took longer to show up in the telecom industry because it is harder to do. This is actually a little counterintuitive to me because in many ways TDM is easier to understand than FDM, but FDM can be implemented with all-analog electronics fairly easily while TDM is hard to do without digital electronics and, ultimately, computers. The basic idea of TDM is that the logical channels take turns. The medium is divided into time slots and each logical channel is assigned a time slot, it gets to speak only during that time slot. TDM is very widely used today because most types of communication media can move data faster than the realtime rate of that data. For example, human speech can be digitized and then transmitted in a shorter period of time than the speech originally took. This means that you can take multiple realtime conversations and pack them onto the same wire by buffering each one to temporary memory and then sending them much faster than they originally occurred during rotating timeslots. TDM is basically old-fashioned sharing and can be visualized (and sometimes implemented) as something like passing a talking stick between logical channels.

Why am I explaining the concepts of FDM and TDM in such depth here? Well, mostly because I am at heart a rambler and once I start on something I cant stop. But also because I think lighting control systems are an interesting opportunity to look at the practicalities of muxing in real systems that are expected to be low-cost, high-reliability, and operate over cabling that isnt too demanding to install.

And also because I think it will be helpful in explaining a historically important lighting control scheme: analog multiplexing, or AMX.

AMX192, the most significant form of AMX, was introduced in 1975 (or so, sources are a little vague on this) by Strand Lighting. Strand is a historically very important manufacturer of theatrical lighting, and later became part of Philips where it was influential on architectural lighting as well (along with the rest of Philips lighting, Strand is now part of Signify). In this way, one can argue that there is a direct through-line from Strands AMX to todays Hue smart bulbs. AMX192 supports 192 channels on a single cable, and uses twisted-pair wiring with two pairs terminated in 4-pin XLR connectors. This will all sound very, very familiar to anyone familiar with theatrical lighting today even if they are too young to have ever dealt with AMX, but well get to that in a bit.

What makes AMX192 (and its broader generation of control protocols) very interesting to me is that it employs analog signaling and TDM. Fundamentally, AMX192 is the same as the 0-10v control scheme (although it actually employs 0-5v), but the analog control signal is sent alongside a clock signal and every clock pulse it changes to the value for the next channel. On the demultiplexing or demuxing side, receivers need to pick out the right channel by counting clock pulses and then freeze the analog value of the signal pair to hold it over while the control wiring cycles through the other channels.

One of the sort of neat things about AMX192 is that you can hook up your control wiring to an oscilloscope and, once youve got the triggering set up right, see a very neat visualization of all 192 control channels across your scope going up and down like the faders on your control board. Its a neat and simple system, but was still fairly cutting edge in the 70s due to the complexity of the electronics used to track the clock pulses.

Well take a moment here too to discuss the physical wiring topology of AMX192: as you might guess, AMX192 is set up as a bus system with each dimmer connected to the same two twisted pairs. In the 70s, dimmers were still fairly large devices and so theaters almost exclusively used traditional dimmer rack systems, with all dimmers installed in one central location. So while there was a multidrop bus wiring arrangement, it was mostly contained to the rack backplanes and not really something that users interacted with.

This idea of multi-drop bus wiring, though, might sound familiar if you have read my other posts. Its largely the same electrical scheme as used by RS-485, a pretty ubiquitous standard for low-speed serial buses. AMX192 is analog, but could RS-485 be applied to use digital signaling on a similar wiring topology?

This is not a hypothetical question, the answer is obviously yes, and about ten years after AMX192 Strand introduced a new digital protocol called DMX512. This stands for Digital Multiplexing, 512 channels, and it employs the RS-485 wiring scheme of one twisted pair in a shielded cable terminated with 5-pin XLR connectors. Now, on the 5-pin XLR connector we have two data pins and one shield/common pin. Of course there are two more pins, and this hints at the curiously complicated landscape of DMX512 cabling.

The DMX512 specification requires that 5-pin cables include two twisted pairs, much like AMX192. You have no doubt determined by now that DMX512 is directly based on AMX192 and carries over the same two-twisted-pair cabling, but with the addition of an extra pin for a grounded shield/signal reference common as required by RS-485, which is the physical layer for DMX512. RS-485 uses embedded clocking though, so it does not require a dedicated pair for clock like AMX192 did. This creates the curious situation that a whole twisted pair is required by the spec but has no specified use. Various off-label applications of the second pair exist, often to carry a second universe of an additional 512 channels, but by far the most alternative use of the second pair is to omit it entirely... resulting in 3 pins, and of course this is a rather attractive option since the 3-pin XLR connector is widely used in live production for balanced audio (e.g. from microphones).

You can run DMX512 over microphone cables, in fact, and it will largely work. A lot of cheaper DMX512 equipment comes fitted with 3-pin XLR connectors for this purpose. The problem is that microphone cables dont actually meet the electrical specifications for DMX512/RS-485 (particularly in that they are not twisted), but on the other hand RS-485 is an intentionally very robust physical protocol and so it tends to work fine in a variety of improper environments. So perhaps a short way to put it is that

DMX512 over 3-pin XLR is probably okay for shorter ranges and if you apply some moral flexibility to standards.

Lets talk a bit about the logical protocol employed by DMX512, because its interesting. DMX512 is a continuous broadcast protocol. That is, despite being digital and packetized it operates exactly like AMX192. The lighting controller continuously transmits the values of every channel in a loop. The only real concession to the power of digital networks in the basic DMX512 protocol is variable slot count. That is, not all 512 channels have to be transmitted if they arent all in use. The controller can send an arbitrary number of channels up to 512. Extensions to the DMX protocol employ a flag byte at the beginning of the frame to support types of messages other than the values for sequential channels starting at 1, but these extensions arent as widely used and tend to be a little more manufacturer-specific.

DMX512 has no error correction or even detection; instead it relies on the fact that all values are repeatedly transmitted so any transient bit error should only be in effect for a short period of time. Of course running DMX512 over non-twisted 3-pin XLR cable will increase the number of such transient errors, and in the modern world of more complex fixtures these errors can become much more noticeable as fixtures stutter in movement.

Lets talk a bit about the fixtures. AMX192 was designed as a solution for the controller to send channel values to the dimmer rack. DMX512 was designed for the same application. The same digital technology that enabled DMX512, though, has enabled a number of innovations in theatrical lighting that could all be summed up as distributed, rather than centralized, dimming. Instead of having a dimmer rack backstage or in a side room, where the dimmers are patched to line-level electrical wiring to fixtures, compact digital dimmers (called dimmer packs) can be placed just about anywhere. DMX512 cabling is then daisy-chained in the simplest configurations or active repeaters are used to distribute the DMX512 frames onto multiple wiring runs.

The next logical step from the dimmer pack is building dimming directly into fixtures, and far more than that has happened. A modern moving head fixture, even a relatively low-end one, can have two axes of movement (altitude-azimuth polar coordinates), four channels of dimming (red, green, blue, white), a multi-position filter or gobo wheel, and even one or two effect drive motors. Higher-end fixtures can have more features like motorized zoom and focus, additional filter wheels and motorized effects, cool white/warm white and UV color channels, etc. The point is that one physical fixture can require direct connection to the DMX bus on which it consumes 8 or more channels. That 512 channel limit can sneak up on you real fast, leading to multi-universe configurations where multiple separate DMX512 networks are used to increase channel count.

DMX, then, while cutting-edge in 1986, is a bit lacking today. Strand basically took AMX192 and shoved it into RS-485 to develop DMX512. Could you take DMX512 and shove it into IP? Consider that a cliffhanger! Theres a *lot* more to this topic, particularly because I havent even started on digital architectural control. While DMX512 can be used for architectural lighting control its not really all that common and theres a universe of interesting protocols on the other side of the fence.

[1] Nonetheless, the effect can be noticeable in theatrical lighting even at its small magnitude with halogen bulbs. As a result many theatrical light controllers have a bulb warmer feature where they keep all fixtures at a very low power level instead of turning them off. You can imagine that when mixing incandescent and LED fixtures with

much more noticeable minimum brightness, making sure this is disabled for the LED fixtures can become a headache.

[2] Some may be familiar with the issue of dimmable vs. non-dimmable fluorescent fixtures, and the similar issue that exists with LEDs to a lesser extent. The difference here is actually less in the light than in the power supply, which for fluorescent fixtures and sometimes LED fixtures is usually called the (whether or not it is actually a ballast in the electrical engineering sense, which newer ballasts are usually not). In LED fixtures it is becoming more common to refer to it as a driver, since the prototypical form of an LED light power supply is a constant-current driver... although once again, many LED drivers are actually more complex devices than simple CC drivers, and the term should be viewed as an imprecise one. Dimmable fluorescent and LED drivers mostly use PWM, meaning that they rapidly switch the output on and off to achieve a desired duty cycle. This is slightly more complicated for fluorescent bulbs due to the need to get them warm enough to strike before they emit light, which means that modern dimmable fluorescent ballasts usually include programmed start. This basically means that theyre running software that detects the state of the lamp based on current consumption and provides striking power if necessary. This is all sort of complicated which is why the dimmable vs. non-dimmable issue is a big one for CFLs and cheaper LED bulbs: in these types of light bulbs the power supply is a large portion of the total cost and simpler non-dimmable ballasts and drivers keep the product price down. Its a much smaller issue in architectural lighting where the type of ballast is specified up front and the ballast is a separate component from the bulb, meaning that its price is a little less important of an issue.